

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2005, 14:45:41 ; Search time 122 Seconds  
(without alignments)  
88.765 Million cell updates/sec

Title: US-10-088-417A-1

Perfect score: 132

Sequence: 1 KIAALKQKIASLKQIEDALEYENDALEQ 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 794064

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	132	100.0	28	4	AAB74343 Peptide S
2	132	100.0	28	4	AAB74351 Peptide S
3	132	100.0	28	8	ADM41436 Self-asee
4	129	97.7	28	4	AAB74341 Peptide S
5	129	97.7	28	4	AAB74342 Peptide S
6	85	64.4	28	8	ADM41438 Self-asee
7	82	62.1	28	4	AAB74345 Peptide S
8	80	60.6	27	8	ADM41437 Self-asee
9	76	57.6	28	4	AAB74348 Peptide S
10	76	57.6	28	4	AAB74356 Peptide S
11	76	57.6	28	4	AAB74352 Peptide S
12	75	56.8	28	4	AAB74346 Peptide S
13	75	56.8	28	4	AAB74347 Peptide S
14	74	56.1	28	4	AAB74349 Peptide S
15	70	53.0	28	4	AAB74355 Peptide u
16	70	53.0	28	4	AAB74350 Peptide S
17	57	43.2	28	7	ABR84737 DE novo d
18	54	40.9	28	7	ABR84736 DE novo d
19	52	39.4	17	4	AAB74344 Peptide C
20	52	39.4	24	5	AAB18799 Strp pep
21	52	39.4	24	5	AAB18797 Strp pep
22	52	39.4	25	3	AAY95958 Coiled-co
23	52	39.4	25	3	AAB08381 Peptide u
24	52	39.4	25	3	AAB08382 Peptide u
25	52	39.4	27	5	AAB18788 CVX0270 p

26	51	38.6	28	4	AAB50879 Integrin
27	51	38.6	28	4	AAB59138 Alpha-hel
28	50	37.9	23	3	AAB08388 Peptide u
29	50	37.9	23	6	ADA00699 Protein m
30	46	34.8	18	4	AAB74340 Peptide C
31	46	34.8	24	2	AAB71601 Lactam br
32	46	34.8	24	3	AAB21683 Coiled co
33	46	34.8	24	7	ADD93917 Coiled co
34	46	34.8	28	2	AAR31979 SSP4 poly
35	45	34.1	24	2	AAW11604 Generic i
36	45	34.1	28	7	ABR84735 DE novo d
37	44	33.3	14	2	AAW71415 Peptide f
38	44	33.3	15	2	AAW71416 Peptide f
39	44	33.3	15	2	AAW80531 Beta-shee
40	44	33.3	16	2	AAW80546 Peptide h
41	44	33.3	18	2	AAW71424 Peptide h
42	44	33.3	21	4	AG65530 Antimicro
43	44	33.3	24	3	AAB21684 Coiled co
44	44	33.3	24	7	ADD93918 Coiled co
45	41	31.1	21	5	ABG78944 Human bre

ALIGNMENTS

RESULT 1  
AAB74343  
ID AAB74343 standard; peptide; 28 AA.

AC AAB74343;

DT 02-JUL-2001 (first entry)

DE Peptide SAF-plC.

KW Atomic Force Microscopy; AFM.

OS Unidentified.

PN WO200121646-A1.

PD 29-MAR-2001.

PF 18-SEP-2000; 2000WO-GB003576.

PR 17-SEP-1999; 99GB-00022013.

PA (UNSU-) UNIV SUSSEX.

PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;

WPI, 2001-335468/35.

PT New protein structures with peptide monomer units, useful in Atomic Force Microscopy, purifying biological fluids, promoting tissue repair and tissue engineering, or constructing nanoscale molecular sieves.

PS Claim 16; Page 26; 45pp; English.

CC The present invention relates to a protein structure with several first peptide monomer units arranged in a first strand and several second peptide monomer units arranged in a second strand. The structure is useful in Atomic Force Microscopy and a number of other applications. The present sequence is the peptide monomer unit SAF-plC

SQ Sequence 28 AA;

Query Match 100.0%; Score 132; DB 4; Length 28;

Best Local Similarity 100.0%; Pred. No. 2.7e-11;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQIEDALEYENDALEQ 28

|||||

DB 1 KIAALKQKIASLKQEIADALEYENDALEQ 28

RESULT 2

AAB74351

ID AAB74351 standard; peptide; 28 AA.

AC AAB74351;

DT 02-JUL-2001 (first entry)

DE Peptide SAF-pl.

KW Atomic Force Microscopy; AFM.

XX Unidentified.

OS WO200121646-A1.

PN 29-MAR-2001.

XX 18-SEP-2000; 2000WO-GB003576.

PR 17-SEP-1999; 99GB-00022013.

XX (UYSU-) UNIV SUSSEX.

PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;

XX WPI; 2001-335468/35.

XX New protein structures with peptide monomer units, useful in Atomic Force Microscopy, purifying biological fluids, promoting tissue repair and tissue engineering, or constructing nanoscale molecular sieves.

PS Claim 16; Page 26; 45pp; English.

XX The present invention relates to a protein structure with several first peptide monomer units arranged in a first strand and several second peptide monomer units arranged in a second strand. The structure is useful in Atomic Force Microscopy and a number of other applications. The present sequence is the peptide monomer unit SAF-pl

XX Sequence 28 AA;

Query Match 100.0%; Score 132; DB 4; Length 28;

Best Local Similarity 100.0%; Pred. No. 2.7e-11;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEIADALEYENDALEQ 28

DB 1 KIAALKQKIASLKQEIADALEYENDALEQ 28

RESULT 3

ADW41436

ID ADW41436 standard; peptide; 28 AA.

AC ADW41436;

DT 03-JUN-2004 (first entry)

DE Self-assembling peptide fibre SAF-pl.

XX Fibre-shaping peptide; self-assembling peptide.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

XX /note= "N-terminal NH3 moiety"

PN WO2004022584-A1.

XX 18-MAR-2004.

XX 08-SEP-2003; 2003WO-GB003900.

XX 06-SEP-2002; 2002GB-00020805.

XX (UYSU-) UNIV SUSSEX.

XX Woolfson D, Ryadnov MG;

XX WPI; 2004-248444/23.

XX Novel fiber-shaping peptide comprising hub and several peptide monomer units, useful for producing protein structure useful in purification of biological fluids and in surface engineering procedures.

XX Example 30; Page 22; 37pp; English.

XX The present sequence is that of self-assembling peptide fibre (SAF) SAF-pl. The invention relates to fibre-shaping (fish) peptides ADM41434-ADM41435 that are capable of interacting with SAFs to form protein structures. The fish peptides allow morphological changes (branches, splits, kinks and bends) to be made to protein fibres comprising SAFs. By incorporating such morphological changes in the protein structures, it is possible to generate a variety of (nanoscale) protein structures, such as assemblies in general, including matrix, filter, network, grid and scaffold structures. Use of the protein structures in the purification of biological fluids, for assembling cells for cell and tissue engineering, and in surface engineering procedures is claimed.

XX Sequence 28 AA;

Query Match 100.0%; Score 132; DB 8; Length 28;

Best Local Similarity 100.0%; Pred. No. 2.7e-11;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEIADALEYENDALEQ 28

DB 1 KIAALKQKIASLKQEIADALEYENDALEQ 28

RESULT 4

AAB74341

ID AAB74341 standard; peptide; 28 AA.

AC AAB74341;

DT 02-JUL-2001 (first entry)

DE Peptide SAF-plA.

XX Atomic Force Microscopy; AFM.

XX Unidentified.

XX WO200121646-A1.

XX 29-MAR-2001.

XX 18-SEP-2000; 2000WO-GB003576.

XX 17-SEP-1999; 99GB-00022013.

XX (UYSU-) UNIV SUSSEX.

XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;

XX WPI; 2001-335468/35.

XX New protein structures with peptide monomer units, useful in Atomic Force Microscopy, purifying biological fluids, promoting tissue repair and tissue engineering, or constructing nanoscale molecular sieves.

XX Claim 16; Page 26; 45pp; English.  
 XX  
 CC The present invention relates to a protein structure with several first  
 CC peptide monomer units arranged in a first strand and several second  
 CC peptide monomer units arranged in a second strand. The structure is  
 CC useful in Atomic Force Microscopy and a number of other applications. The  
 CC present sequence is the peptide monomer unit SAF-p1a  
 XX  
 SQ Sequence 28 AA;

Query Match 97.7%; Score 129; DB 4; Length 28;  
 Best Local Similarity 96.4%; Pred. No. 7e-11;  
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KIAALKQKIASLKQEIDALEYENDALEQ 28  
 DB 1 KIAALKQKIASLKQEIDALEYENDALEQ 28

RESULT 5  
 AAB74342  
 ID AAB74342 standard; peptide; 28 AA.  
 XX  
 AC AAB74342;  
 XX  
 DT 02-JUL-2001 (first entry)  
 XX  
 DE Peptide SAF-p1B.  
 XX  
 KW Atomic Force Microscopy; AFM.  
 XX  
 OS Unidentified.  
 XX  
 FN WO200121646-A1.  
 XX  
 PD 29-MAR-2001.  
 XX  
 PF 18-SEP-2000; 2000WO-GB003576.  
 XX  
 PR 17-SEP-1999; 99GB-00022013.  
 XX  
 PA (UYSU-) UNIV SUSSEX.  
 XX  
 PI Woolfson DN, Waleshaw J, Pandya MJ, Colyer J;  
 XX  
 DR WPI; 2001-335468/35.  
 XX  
 PT New protein structures with peptide monomer units, useful in Atomic Force  
 PT Microscopy, purifying biological fluids, promoting tissue repair and  
 PT tissue engineering, or constructing nanoscale molecular sieves.  
 XX  
 PS Disclosure; Page 9; 45pp; English.  
 XX  
 CC The present invention relates to a protein structure with several first  
 CC peptide monomer units arranged in a first strand and several second  
 CC peptide monomer units arranged in a second strand. The structure is  
 CC useful in Atomic Force Microscopy and a number of other applications. The  
 CC present sequence is the peptide monomer unit SAF-plB  
 XX  
 SQ Sequence 28 AA;

Query Match 97.7%; Score 129; DB 4; Length 28;  
 Best Local Similarity 96.4%; Pred. No. 7e-11;  
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KIAALKQKIASLKQEIDALEYENDALEQ 28  
 DB 1 KIAALKQKIASLKQEIDALEYENDALEQ 28

RESULT 6  
 ADM41438

ID ADM41438 standard; peptide; 28 AA.  
 XX  
 AC ADM41438;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Self-assembling peptide fibre SAF-p2a.  
 XX  
 KW Fibre-shaping peptide; self-assembling peptide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal NH3 moiety"  
 XX  
 PN WO2004022584-A1.  
 XX  
 PD 18-MAR-2004.  
 XX  
 PF 08-SEP-2003; 2003WO-GB003900.  
 XX  
 PR 06-SEP-2002; 2002GB-00020805.  
 XX  
 PA (UYSU-) UNIV SUSSEX.  
 XX  
 PI Woolfson D, Ryadnov MG;  
 XX  
 DR WPI; 2004-248444/23.  
 XX  
 PT Novel fiber-shaping peptide comprising hub and several peptide monomer  
 PT units, useful for producing protein structure useful in purification of  
 PT biological fluids and in surface engineering procedures.  
 XX  
 PS Example 36; Page 22; 37pp; English.  
 XX  
 CC The present sequence is that of self-assembling peptide fibre (SAF) SAF-  
 CC p2a. The invention relates to fibre-shaping (Fish) peptides ADM41434-  
 CC ADM41435 that are capable of interacting with SAFs to form protein  
 CC structures. The Fish peptides allow morphological changes (branches,  
 CC splits, kinks and bends) to be made to protein fibres comprising SAFs. By  
 CC incorporating such morphological changes in the protein fibres, it is  
 CC possible to generate a variety of (nanoscale) protein structures, such as  
 CC assemblies in general, including matrix, filter, network, grid and  
 CC scaffold structures. Use of the protein structures in the purification of  
 CC biological fluids, for assembling cells for cell and tissue engineering,  
 CC and in surface engineering procedures is claimed.  
 XX  
 SQ Sequence 28 AA;  
 Query Match 64.4%; Score 85; DB 8; Length 28;  
 Best Local Similarity 75.0%; Pred. No. 9.5e-05;  
 Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 KIAALKQKIASLKQEIDALEYENDALEQ 28  
 DB 1 KIRRLKQKNARLKQEIADALEYEALEQ 28  
 RESULT 7  
 AAB74345  
 ID AAB74345 standard; peptide; 28 AA.  
 XX  
 AC AAB74345;  
 XX  
 DT 02-JUL-2001 (first entry)  
 XX  
 DE Peptide SAF-p2A.  
 XX  
 KW Atomic Force Microscopy; AFM.  
 XX  
 OS Unidentified.  
 XX

PN WO200121646-A1.  
XX 29-MAR-2001.  
XX 18-SEP-2000; 2000WO-GB003576.  
XX 17-SEP-1999; 99GB-00022013.  
XX (UYSU-) UNIV SUSSEX.  
XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;  
XX WPI; 2001-335468/35.  
XX New protein structures with peptide monomer units, useful in Atomic Force  
PT Microscopy, purifying biological fluids, promoting tissue repair and  
PT tissue engineering, or constructing nanoscale molecular sieves.  
XX Disclosure; Page 9; 45pp; English.  
XX The present invention relates to a protein structure with several first  
CC peptide monomer units arranged in a first strand and several second  
CC peptide monomer units arranged in a second strand. The structure is  
CC useful in Atomic Force Microscopy and a number of other applications. The  
CC present sequence is the peptide monomer unit SAF-p2A  
XX Sequence 28 AA;  
SQ Query Match 62.1%; Score 82; DB 4; Length 28;  
Best Local Similarity 75.0%; Pred. No. 0.00025;  
Matches 21; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 1 KIAALKQKIASLKQEIADALEYENDALEQ 28  
DB 1 KISALKQKIASLKQEIADALEYENDALEQ 28  
RESULT 8  
ADM41437  
ID ADM41437 standard; peptide; 27 AA.  
XX ADM41437;  
XX 03-JUN-2004 (first entry)  
XX Self-assembling peptide fibre (SAF).  
XX Fibre-shaping peptide; self-assembling peptide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal NH3 moiety"  
FT FT  
XX WO2004022584-A1.  
XX 18-MAR-2004.  
XX 08-SEP-2003; 2003WO-GB003900.  
XX 06-SEP-2002; 2002GB-00020805.  
XX (UYSU-) UNIV SUSSEX.  
XX Woolfson D, Ryadnov MG;  
XX WPI; 2004-248444/23.  
XX Novel fiber-shaping peptide comprising hub and several peptide monomer  
PT units, useful for producing protein structure useful in purification of  
PT biological fluids and in surface engineering procedures.  
XX

PS Example 30; Page 22; 37pp; English.  
XX The present sequence is that of a self-assembling peptide fibre (SAF).  
CC The invention relates to fibre-shaping (FISH) peptides ADM41434-ADM41435  
CC that are capable of interacting with SAFs to form protein structures. The  
CC FISH peptides allow morphological changes (branches, splits, kinks and  
CC bends) to be made to protein fibres comprising SAFs. By incorporating  
CC such morphological changes in the protein fibres, it is possible to  
CC generate a variety of (nanoscale) protein structures, such as assemblies  
CC in general, including matrix, filter, network, grid and scaffold  
CC structures. Use of the protein structures in the purification of  
CC biological fluids, for assembling cells for cell and tissue engineering,  
CC and in surface engineering procedures is claimed.  
XX Sequence 27 AA;  
SQ Query Match 50.6%; Score 80; DB 8; Length 27;  
Best Local Similarity 74.1%; Pred. No. 0.00045;  
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 2 IAALKQKIASLKQEIADALEYENDALEQ 28  
DB 1 IRRLLKQKIASLKQEIADALEYENDALEQ 27  
RESULT 9  
AAB74348  
ID AAB74348 standard; peptide; 28 AA.  
XX AAB74348;  
XX 02-JUL-2001 (first entry)  
XX Peptide SAF-p2D.  
XX Atomic Force Microscopy; AFM.  
XX Unidentified.  
XX WO200121646-A1.  
XX 29-MAR-2001.  
XX 18-SEP-2000; 2000WO-GB003576.  
XX 17-SEP-1999; 99GB-00022013.  
XX (UYSU-) UNIV SUSSEX.  
XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;  
XX WPI; 2001-335468/35.  
XX New protein structures with peptide monomer units, useful in Atomic Force  
PT Microscopy, purifying biological fluids, promoting tissue repair and  
PT tissue engineering, or constructing nanoscale molecular sieves.  
XX Disclosure; Page 9; 45pp; English.  
XX The present invention relates to a protein structure with several first  
CC peptide monomer units arranged in a first strand and several second  
CC peptide monomer units arranged in a second strand. The structure is  
CC useful in Atomic Force Microscopy and a number of other applications. The  
CC present sequence is the peptide monomer unit SAF-p2D  
XX Sequence 28 AA;  
SQ Query Match 57.6%; Score 76; DB 4; Length 28;  
Best Local Similarity 71.4%; Pred. No. 0.0017;  
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 KIAALKQKIASLKQEIADALEYENDALEQ 28  
DB 1 KISALKQKIASLKQEIADALEYENDALEQ 28

```
Db      1 KIRALKAKNAHLKQEIATIALEQIEIAALEQ 28

RESULT 10
AAB74356
ID AAB74356 standard; peptide; 28 AA.
XX
XX AC AAB74356;
XX
XX DT 02-JUL-2001 (first entry)
XX
XX DE Peptide used to form blunt-ended heterodimers.
XX
XX KW Atomic Force Microscopy; AFM.
XX
XX OS Unidentified.
XX
XX PN WO200121646-A1.
XX
XX PD 29-MAR-2001.
XX
XX PF 18-SEP-2000; 2000WO-GB003576.
XX
XX PR 17-SEP-1999; 99GB-00022013.
XX
XX PA (UYSU-) UNIV SUSSEX.
XX
XX PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;
XX
XX DR WPI; 2001-335468/35.
XX
XX PT New protein structures with peptide monomer units, useful in Atomic Force
XX
XX PT Microscopy, purifying biological fluids, promoting tissue repair and
XX
XX PT tissue engineering, or constructing nanoscale molecular sieves.
XX
XX PS Disclosure; Fig 8; 45pp; English.
XX
XX OS The present invention relates to a protein structure with several first
XX
XX CC peptide monomer units arranged in a first strand and several second
XX
XX CC peptide monomer units arranged in a second strand. The structure is
XX
XX CC useful in Atomic Force Microscopy and a number of other applications. The
XX
XX CC present sequence is a peptide used to form blunt-ended heterodimers
XX
XX SQ Sequence 28 AA;

Query Match 57.6%; Score 76; DB 4; Length 28;
Best Local Similarity 71.4%; Pred. No. 0.0017;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 KIRALKKQKIASLKQEIADALEYENDALEQ 28
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 KIRALKKNAHLKQEIATIALEQIEIAALEQ 28

RESULT 11
AAB74352
ID AAB74352 standard; peptide; 28 AA.
XX
XX AC AAB74352;
XX
XX DT 02-JUL-2001 (first entry)
XX
XX DE Peptide SAF-p2.
XX
XX KW Atomic Force Microscopy; AFM.
XX
XX OS Unidentified.
XX
XX PN WO200121646-A1.
XX
XX PD 29-MAR-2001.
XX
XX PF 18-SEP-2000; 2000WO-GB003576.

Query Match 57.6%; Score 76; DB 4; Length 28;
Best Local Similarity 71.4%; Pred. No. 0.0017;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 KIRALKKQKIASLKQEIADALEYENDALEQ 28
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 KIRALKKNAHLKQEIATIALEQIEIAALEQ 28

RESULT 12
AAB74346
ID AAB74346 standard; peptide; 28 AA.
XX
XX AC AAB74346;
XX
XX DT 02-JUL-2001 (first entry)
XX
XX DE Peptide SAF-p2B.
XX
XX KW Atomic Force Microscopy; AFM.
XX
XX OS Unidentified.
XX
XX PN WO200121646-A1.
XX
XX PD 29-MAR-2001.
XX
XX PF 18-SEP-2000; 2000WO-GB003576.
XX
XX PR 17-SEP-1999; 99GB-00022013.
XX
XX PA (UYSU-) UNIV SUSSEX.
XX
XX PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;
XX
XX DR WPI; 2001-335468/35.
XX
XX PT New protein structures with peptide monomer units, useful in Atomic Force
XX
XX PT Microscopy, purifying biological fluids, promoting tissue repair and
XX
XX PT tissue engineering, or constructing nanoscale molecular sieves.
XX
XX PS Disclosure; Page 9; 45pp; English.
XX
XX OS The present invention relates to a protein structure with several first
XX
XX CC peptide monomer units arranged in a first strand and several second
XX
XX CC peptide monomer units arranged in a second strand. The structure is
XX
XX CC useful in Atomic Force Microscopy and a number of other applications. The
XX
XX CC present sequence is the peptide monomer unit SAF-p2B
XX
XX SQ Sequence 28 AA;
```

```

Query Match      56.8%; Score 75; DB 4; Length 28;
Best Local Similarity 71.4%; Pred. No. 0.0024;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEIADALEVENDALEQ 28
DB 1 KIRALKKKNVHLKQEIADALEVENDALEQ 28

RESULT 13
AAB74347
ID AAB74347 standard; peptide; 28 AA.
XX
AC AAB74347;
XX
DT 02-JUL-2001 (first entry)
XX
DE Peptide SAF-p2C.
XX
KW Atomic Force Microscopy; AFM.
XX
OS Unidentified.
XX
PN WO200121646-A1.
XX
PD 29-MAR-2001.
XX
PF 18-SEP-2000; 2000WO-GB003576.
XX
PR 17-SEP-1999; 99GB-00022013.
XX
PA (UYSU-) UNIV SUSSEX.
XX
PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;
XX
WPI; 2001-335468/35.
XX
New protein structures with peptide monomer units, useful in Atomic Force
PT Microscopy, purifying biological fluids, promoting tissue repair and
PT tissue engineering, or constructing nanoscale molecular sieves.
XX
PS Disclosure; Page 9; 45pp; English.
XX
CC The present invention relates to a protein structure with several first
CC peptide monomer units arranged in a first strand and several second
CC peptide monomer units arranged in a second strand. The structure is
CC useful in Atomic Force Microscopy and a number of other applications. The
CC present sequence is the peptide monomer unit SAF-p2C
XX
SQ Sequence 28 AA;

Query Match      56.8%; Score 75; DB 4; Length 28;
Best Local Similarity 71.4%; Pred. No. 0.0024;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEIADALEVENDALEQ 28
DB 1 KIRALKKKNVHLKQEIADALEVENDALEQ 28

RESULT 14
AAB74349
ID AAB74349 standard; peptide; 28 AA.
XX
AC AAB74349;
XX
DT 02-JUL-2001 (first entry)
XX
DE Peptide SAF-p2E.
XX
KW Atomic Force Microscopy; AFM.
XX

```

```

OS Unidentified.
XX
PN WO200121646-A1.
XX
PD 29-MAR-2001.
XX
PF 18-SEP-2000; 2000WO-GB003576.
XX
PR 17-SEP-1999; 99GB-00022013.
XX
PA (UYSU-) UNIV SUSSEX.
XX
PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;
XX
WPI; 2001-335468/35.
XX
New protein structures with peptide monomer units, useful in Atomic Force
PT Microscopy, purifying biological fluids, promoting tissue repair and
PT tissue engineering, or constructing nanoscale molecular sieves.
XX
PS Disclosure; Page 9; 45pp; English.
XX
CC The present invention relates to a protein structure with several first
CC peptide monomer units arranged in a first strand and several second
CC peptide monomer units arranged in a second strand. The structure is
CC useful in Atomic Force Microscopy and a number of other applications. The
CC present sequence is the peptide monomer unit SAF-p2E
XX
SQ Sequence 28 AA;

Query Match      56.1%; Score 74; DB 4; Length 28;
Best Local Similarity 71.4%; Pred. No. 0.0033;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEIADALEVENDALEQ 28
DB 1 KIRALKKKNVHLKQEIADALEVENDALEQ 28

RESULT 15
AAB74355
ID AAB74355 standard; peptide; 28 AA.
XX
AC AAB74355;
XX
DT 02-JUL-2001 (first entry)
XX
DE Peptide used to form blunt-ended heterodimers.
XX
KW Atomic Force Microscopy; AFM.
XX
OS Unidentified.
XX
PN WO200121646-A1.
XX
PD 29-MAR-2001.
XX
PF 18-SEP-2000; 2000WO-GB003576.
XX
PR 17-SEP-1999; 99GB-00022013.
XX
PA (UYSU-) UNIV SUSSEX.
XX
PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;
XX
WPI; 2001-335468/35.
XX
New protein structures with peptide monomer units, useful in Atomic Force
PT Microscopy, purifying biological fluids, promoting tissue repair and
PT tissue engineering, or constructing nanoscale molecular sieves.
XX
PS Disclosure; Fig 8; 45pp; English.
XX

```

CC The present invention relates to a protein structure with several first  
 CC peptide monomer units arranged in a first strand and several second  
 CC peptide monomer units arranged in a second strand. The structure is  
 CC useful in Atomic Force Microscopy and a number of other applications. The  
 CC present sequence is a peptide used to form blunt-ended heterodimers  
 XX

SQ Sequence 28 AA;

Query Match 53.0%; Score 70; DB 4; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 EIDALEYENDALEQ 28  
 |||||  
 Db 1 EIDALEYENDALEQ 14

Search completed: April 27, 2005, 15:02:15  
 Job time : 124 secs

**BEST AVAILABLE COPY**

**THIS PAGE LEFT BLANK**



Result No.	Score	Query Match	Length	DB	ID	Description
1	57	43.2	28	14	US-10-360-053-23	Sequence 23, Appl
2	54	40.9	28	14	US-10-360-053-22	Sequence 22, Appl
3	52	39.4	27	10	US-09-882-774-7	Sequence 7, Appl
4	51	38.6	28	9	US-09-320-907B-6	Sequence 6, Appl
5	51	38.6	28	15	US-10-447-292-6	Sequence 6, Appl
6	50	37.9	23	10	US-09-259-658-59	Sequence 59, Appl
7	46	34.8	24	14	US-10-338-083-19	Sequence 19, Appl
8	46	34.8	24	16	US-10-611-399-19	Sequence 19, Appl
9	46	34.8	24	17	US-10-794-751-19	Sequence 19, Appl
10	45	34.1	28	14	US-10-360-053-21	Sequence 21, Appl
11	44	33.3	24	14	US-10-338-083-20	Sequence 20, Appl
12	44	33.3	24	16	US-10-611-399-20	Sequence 20, Appl
13	44	33.3	24	17	US-10-794-751-20	Sequence 20, Appl



APPLICANT: Craig  
APPLICANT: Maschio  
APPLICANT: Mezna  
TITLE OF INVENTION: Compositions And Methods For Monitoring The  
Modification State Of A Pair Of Polypeptides  
FILE REFERENCE: colyer 456/79245  
CURRENT APPLICATION NUMBER: US/09/259,658  
CURRENT FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 59  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptide used to assay geranylgeranyl transferase  
OTHER INFORMATION: activity.  
US-09-259-658-59

Query Match 37.9%; Score 50; DB 10; Length 23;  
Best Local Similarity 75.0%; Pred. No. 10;  
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KIAALKQKIASLKQEI 16  
||| ||||| |||||  
Db 7 KIAQLKQKNACLKQKI 22

## RESULT 7

US-10-338-083-19  
Sequence 19, Application US/10338083  
Publication No. US20030166559A1  
GENERAL INFORMATION:

APPLICANT: Desjarlais, John R.  
APPLICANT: Tansey, Malu Lourdes G.  
APPLICANT: Dahiyat, Bassil I.  
TITLE OF INVENTION: Dominant Negative Proteins and Methods Thereof  
FILE REFERENCE: A-71273-2  
CURRENT APPLICATION NUMBER: US/10/338,083  
CURRENT FILING DATE: 2003-01-16  
PRIOR APPLICATION NUMBER: US 60/345,805  
PRIOR FILING DATE: 2002-01-04  
PRIOR APPLICATION NUMBER: US 60/373,453  
PRIOR FILING DATE: 2002-04-17  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 19  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: coiled-coil motif  
US-10-338-083-19

Query Match 34.8%; Score 46; DB 14; Length 24;  
Best Local Similarity 41.7%; Pred. No. 34;  
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 3 AALKQKIASLKQEI DALEYENDAL 26  
||| : : : : : : : : : : : :  
Db 1 AALESEVSALESEVAAL 24

## RESULT 8

US-10-611-399-19  
Sequence 19, Application US/10611399  
Publication No. US20040170602A1  
GENERAL INFORMATION:

APPLICANT: Desjarlais, John R.  
APPLICANT: Tansey, Malu Lourdes G.  
APPLICANT: Dahiyat, Bassil I.  
TITLE OF INVENTION: DOMINANT NEGATIVE PROTEINS AND METHODS THEREOF

FILE REFERENCE: A-71273-3  
CURRENT APPLICATION NUMBER: US/10/611,399  
CURRENT FILING DATE: 2003-07-01  
PRIOR APPLICATION NUMBER: US 10/338,083  
PRIOR FILING DATE: 2003-01-06  
PRIOR APPLICATION NUMBER: US 60/345,805  
PRIOR FILING DATE: 2002-01-04  
PRIOR APPLICATION NUMBER: US 60/373,453  
PRIOR FILING DATE: 2002-04-17  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 19  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: coiled-coil motif  
US-10-611-399-19

Query Match 34.8%; Score 46; DB 16; Length 24;  
Best Local Similarity 41.7%; Pred. No. 34;  
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 3 AALKQKIASLKQEI DALEYENDAL 26  
||| : : : : : : : : : : : :  
Db 1 AALESEVSALESEVAAL 24

## RESULT 9

US-10-794-751-19  
Sequence 19, Application US/10794751  
Publication No. US20050048626A1  
GENERAL INFORMATION:

APPLICANT: Desjarlais, John R.  
APPLICANT: Thomason, Adam Read  
APPLICANT: Zhukovsky, Eugene Alexander  
TITLE OF INVENTION: BAFF VARIANTS AND METHODS THEREOF  
FILE REFERENCE: A-72175-1  
CURRENT APPLICATION NUMBER: US/10/794,751  
CURRENT FILING DATE: 2004-03-05  
PRIOR APPLICATION NUMBER: US 10/338,083  
PRIOR FILING DATE: 2003-01-06  
PRIOR APPLICATION NUMBER: US 60/452,707  
PRIOR FILING DATE: 2003-03-07  
PRIOR APPLICATION NUMBER: US 60/482,081  
PRIOR FILING DATE: 2003-06-23  
PRIOR APPLICATION NUMBER: US 60/523,880  
PRIOR FILING DATE: 2003-11-20  
PRIOR APPLICATION NUMBER: US 60/528,104  
PRIOR FILING DATE: 2003-12-08  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 19  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: coiled-coil motif  
US-10-794-751-19

Query Match 34.8%; Score 46; DB 17; Length 24;  
Best Local Similarity 41.7%; Pred. No. 34;  
Matches 10; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 3 AALKQKIASLKQEI DALEYENDAL 26  
||| : : : : : : : : : : : :  
Db 1 AALESEVSALESEVAAL 24

## RESULT 10

US-10-360-053-21  
Sequence 21, Application US/10360053  
Publication No. US20030170230A1

```

; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Utenthal, Lars O
; APPLICANT: Nielsen, Rasmus W
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibody
; TITLE OF INVENTION: Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-21

Query Match 34.1%; Score 45; DB 14; Length 28;
Best Local Similarity 56.2%; Pred. No. 55;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEI 16
Db 12 EIAAIKDKIAAIKEYI 27

RESULT 11
US-10-338-083-20
; Sequence 20, Application US/10338083
; Publication No. US20030166559A1
; GENERAL INFORMATION:
; APPLICANT: Desjarlais, John R.
; APPLICANT: Tansey, Malu Lourdes G.
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: Dominant Negative Proteins and Methods Thereof
; FILE REFERENCE: A-71273-2
; CURRENT APPLICATION NUMBER: US/10/338,083
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/345,805
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/373,453
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: coiled-coil motif
US-10-338-083-20

Query Match 33.3%; Score 44; DB 14; Length 24;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEI 18
Db 7 KLSAVKSKLASVSKLAA 24

RESULT 12
US-10-611-399-20
; Sequence 20, Application US/10611399
; Publication No. US20040170602A1
; GENERAL INFORMATION:
; APPLICANT: Desjarlais, John R.
; APPLICANT: Tansey, Malu Lourdes G.

```

```

; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: DOMINANT NEGATIVE PROTEINS AND METHODS THEREOF
; FILE REFERENCE: A-71273-3
; CURRENT APPLICATION NUMBER: US/10/611,399
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 10/338,083
; PRIOR FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US 60/345,805
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/373,453
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: coiled-coil motif
US-10-611-399-20

Query Match 33.3%; Score 44; DB 16; Length 24;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEI 18
Db 7 KLSAVKSKLASVSKLAA 24

RESULT 13
US-10-794-751-20
; Sequence 20, Application US/10794751
; Publication No. US20050048626A1
; GENERAL INFORMATION:
; APPLICANT: Desjarlais, John R.
; APPLICANT: Thomason, Adam Read
; APPLICANT: Zhukovsky, Eugene Alexander
; TITLE OF INVENTION: BAFF VARIANTS AND METHODS THEREOF
; FILE REFERENCE: A-72175-1
; CURRENT APPLICATION NUMBER: US/10/794,751
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 10/338,083
; PRIOR FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US 60/452,707
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: US 60/482,081
; PRIOR FILING DATE: 2003-06-23
; PRIOR APPLICATION NUMBER: US 60/523,880
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: US 60/528,104
; PRIOR FILING DATE: 2003-12-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: coiled-coil motif
US-10-794-751-20

Query Match 33.3%; Score 44; DB 17; Length 24;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQEI 18
Db 7 KLSAVKSKLASVSKLAA 24

RESULT 14
US-10-630-926-7

```

```
; Sequence 7, Application US/10630926
; Publication NO. US20040194160A1
; GENERAL INFORMATION:
; APPLICANT: RICCARDI, Carlo
; TITLE OF INVENTION: INTRACELLULAR MODULATORS OF APOPTIC CELL
; DEATH PATHWAYS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/630,926
; FILING DATE: 31-Jul-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/403,861A
; FILING DATE: 11-Feb-2000
; APPLICATION NUMBER: PCT/EP98/02490
; FILING DATE: 27-APR-1998
; APPLICATION NUMBER: EP 97107033.9
; FILING DATE: 28-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: RICCARDI=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-630-926-7

Query Match 33.3%; Score 44; DB 16; Length 26;
Best Local Similarity 43.5%; Pred. No. 68;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 5 LKQKIASLKEIDALEYENDALE 27
Db 1 LKEQIKELIEKNSQLEQENDLLK 23

RESULT 15
US-09-834-759-534
; Sequence 534, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 534
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-534
```

```
Query Match 31.1%; Score 41; DB 9; Length 21;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LKQKIASLKEIDALEY 22
Db 4 LKKEIAMLKLEIATLKHQ 21
```

```
Search completed: April 27, 2005, 15:22:47
Job time : 95.5 secs
```

**BEST AVAILABLE COPY**

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2005, 14:58:17 ; Search time 29.5 Seconds  
(without alignments)  
70.853 Million cell updates/sec

Title: US-10-088-417A-1  
Perfect score: 132  
Sequence: 1 KIAALKQKIASLKQIDALEYENDALEQ 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 228236

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	39.4	25	4	US-09-491-614B-25
2	46	34.8	24	2	US-08-491-527A-12
3	46	34.8	28	1	US-08-182-175A-1
4	46	34.8	28	5	PCT-US92-06412-1
5	45	34.1	24	2	US-08-491-527A-11
6	44	33.3	26	4	US-09-403-861A-7
7	41	31.1	21	4	US-09-834-759-534
8	40	30.3	28	1	US-08-182-175A-3
9	40	30.3	28	1	US-08-474-633A-70
10	40	30.3	28	4	US-08-823-771-70
11	40	30.3	28	5	PCT-US92-06412-3
12	39	29.5	25	3	US-08-737-629-8
13	39	29.5	27	1	US-08-446-922-9
14	39	29.5	27	2	US-08-484-624A-24
15	39	29.5	27	2	US-08-477-733B-24
16	39	29.5	27	3	US-09-088-913A-24
17	39	29.5	27	3	US-08-737-629-1
18	39	29.5	27	3	US-08-769-819-24
19	39	29.5	27	3	US-09-320-424-14
20	39	29.5	27	3	US-08-770-974-24
21	39	29.5	27	3	US-08-770-981-24
22	39	29.5	27	4	US-09-399-106-24
23	39	29.5	27	4	US-09-645-926A-4
24	39	29.5	27	4	US-09-509-802-4
25	39	29.5	27	4	US-09-825-563-14
26	39	29.5	27	4	US-09-852-391-4
27	39	29.5	27	4	US-09-524-100C-12

28	39	29.5	27	4	US-09-976-472A-4	Sequence 4, Appl
29	38	28.8	25	3	US-08-690-011A-2	Sequence 2, Appl
30	38	28.8	25	3	US-09-299-495F-2	Sequence 2, Appl
31	38	28.8	26	2	US-08-690-011A-3	Sequence 3, Appl
32	38	28.8	26	3	US-09-299-495F-3	Sequence 3, Appl
33	37	28.0	26	2	US-08-690-011A-37	Sequence 37, Appl
34	37	28.0	26	3	US-09-299-495F-37	Sequence 37, Appl
35	37	28.0	27	4	US-09-664-945-77	Sequence 77, Appl
36	37	28.0	28	1	US-08-182-175A-2	Sequence 2, Appl
37	37	28.0	28	1	US-08-182-175A-39	Sequence 39, Appl
38	37	28.0	28	1	US-08-182-175A-43	Sequence 43, Appl
39	37	28.0	28	1	US-08-182-175A-47	Sequence 47, Appl
40	37	28.0	28	1	US-08-474-633A-36	Sequence 36, Appl
41	37	28.0	28	1	US-08-474-633A-40	Sequence 40, Appl
42	37	28.0	28	1	US-08-474-633A-56	Sequence 56, Appl
43	37	28.0	28	1	US-08-474-633A-67	Sequence 67, Appl
44	37	28.0	28	1	US-08-944-133-18	Sequence 18, Appl
45	37	28.0	28	1	US-08-944-133-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1

US-09-491-614B-25  
; Sequence 25, Application US/09491614B  
; Patent No. 6828106  
; GENERAL INFORMATION:  
; APPLICANT: Colyer, John  
; TITLE OF INVENTION: Methods and Compositions Using Coiled Binding Partners  
; FILE REFERENCE: 10069/1150  
; CURRENT APPLICATION NUMBER: US/09/491,614B  
; CURRENT FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 09/259,474  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 25  
; LENGTH: 25  
; TYPE: PPT  
; ORGANISM: artificial  
; FEATURE:  
; OTHER INFORMATION: Peptide for forming a coiled coil  
; NAME/KEY: Acetylation  
; LOCATION: (1)..(1)  
; FEATURE:  
; NAME/KEY: Amidation  
; LOCATION: (25)..(25)  
US-09-491-614B-25

Query Match 39.4%; Score 52; DB 4; Length 25;  
Best Local Similarity 60.0%; Pred. No. 0.69; Mismatches 3; Indels 0; Gaps 0;  
Matches 12; Conservative 5

Qy 9 IASLKQIDALEYENDALEQ 28  
|||:|||||  
Db 1 IAALEREIYKLEQENQLEQ 20

RESULT 2

US-08-491-527A-12  
; Sequence 12, Application US/08491527A  
; Patent No. 5824483  
; GENERAL INFORMATION:  
; APPLICANT: Houston, Michael E.  
; APPLICANT: Hodges, Robert S.  
; TITLE OF INVENTION: Conformationally-Restricted Combinatorial  
; TITLE OF INVENTION: Library Composition and Method  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Denlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/743,006
; FILING DATE: 9 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda Axamethy Floyd
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB-1031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4929
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..28
; OTHER INFORMATION: /label= name
; OTHER INFORMATION: /note= "(SSP 4)"
;
US-08-182-175A-1
Query Match 34.8%; Score 46; DB 1; Length 28;
Best Local Similarity 40.0%; Pred. No. 5.5;
Matches 8; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQSIDALE 20
|: ||::: |:|:|
DB 4 KIKALEEKLIKALEKIKALE 23

RESULT 4
PCT-US92-06412-1
; Sequence 1, Application PC/TUS9206412
; GENERAL INFORMATION:
; APPLICANT: Saverio Carl Falco
; APPLICANT: Sharon J. Keeler
; APPLICANT: Janet A. Rice
; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Con
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06412
; FILING DATE: 19920807
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/743,006
; FILING DATE: 9 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda Axamethy Floyd
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: BB-1031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4929
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
;

```



;; TYPE: AMINO ACID  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; FEATURE:  
;; NAME/KEY: Protein  
;; LOCATION: 1..28  
;; OTHER INFORMATION: /label= name  
;; OTHER INFORMATION: /note= "(SSP 4)"  
PCT-US92-06412-1

Query Match 34.8%; Score 46; DB 5; Length 28;  
Best Local Similarity 40.0%; Pred. No. 5.5;  
Matches 8; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KIAALKQKIASLKQEIADLE 20  
Db 4 KLKALEEKLKALEEKLKALE 23

RESULT 5  
US-08-491-527A-11  
; Sequence 11, Application US/08491527A  
; Patent No. 5824483  
; GENERAL INFORMATION:  
; APPLICANT: Houston, Michael E.  
; APPLICANT: Hodges, Robert S.  
; TITLE OF INVENTION: Conformationally-Restricted Combinatorial  
; TITLE OF INVENTION: Library Composition and Method  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/491,527A  
FILING DATE: 16-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,199  
FILING DATE: 15-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/245,507  
FILING DATE: 18-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 7900-0008.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: exemplary generic library peptide,  
INDIVIDUAL ISOLATE: Fig. 3A

US-08-491-527A-11

Query Match 34.1%; Score 45; DB 2; Length 24;

Best Local Similarity 50.0%; Pred. No. 6.3;  
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 KIAALKQKIASLKQEIADLE 22  
Db 1 EIEALKKEIXLXXKIXALEKE 22

RESULT 6  
US-09-403-861A-7  
; Sequence 7, Application US/09403861A  
; Patent No. 6833348  
; GENERAL INFORMATION:  
; APPLICANT: RICCARDI, Carlo  
; TITLE OF INVENTION: INTRACELLULAR MODULATORS OF APOPTIC CELL  
; DEATH PATHWAYS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/403,861A  
FILING DATE: 11-Feb-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP98/02490  
FILING DATE: 27-APR-1998  
APPLICATION NUMBER: EP 97107033.9  
FILING DATE: 28-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: YUN, Allen C.

REGISTRATION NUMBER: 37,971  
REFERENCE/DOCKET NUMBER: RICCARDI=1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-403-861A-7

Query Match 33.3%; Score 44; DB 4; Length 26;  
Best Local Similarity 43.5%; Pred. No. 9.6;  
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 5 LKQKIASLKQEIADLE 27  
Db 1 LKQKIELIKSQEQLDLE 23

RESULT 7  
US-09-834-759-534  
; Sequence 534, Application US/09834759  
; Patent No. 6880197  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Dillon, Devin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Hepler, William T.

APPLICANT: Henderson, Robert A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.470C9  
CURRENT APPLICATION NUMBER: US/09/834,759  
CURRENT FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 547  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 534  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-834-759-534

Query Match 31.1%; Score 41; DB 4; Length 21;  
Best Local Similarity 50.0%; Pred. No. 20;  
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 LKQKIASLKQKIDALEYE 22  
DB 4 LKKEIAMLKLEIATLKHQ 21

RESULT 8  
US-08-182-175A-3  
Sequence 3, Application US/08182175A  
Patent No. 5559223  
GENERAL INFORMATION:  
APPLICANT: Saverio Carl Falco  
APPLICANT: Sharon J. Keeler  
APPLICANT: Janet A. Rice  
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing B  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. du Pont de Nemours and Company  
STREET: 1007 Market Street  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/182,175A  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/743,006  
FILING DATE: 9 August 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Linda Axamethy Floyd  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: BB-1031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 992-4929  
TELEFAX: (302) 892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..28  
OTHER INFORMATION: /label= name  
OTHER INFORMATION: /note= "(SSP 7)4"

US-08-182-175A-3

Query Match 30.3%; Score 40; DB 1; Length 28;  
Best Local Similarity 25.0%; Pred. No. 38;  
Matches 5; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQKIDALE 20  
DB 4 KKKAMEEKKLAMEEKLKAME 23

RESULT 9  
US-08-474-633A-70  
Sequence 70, Application US/08474633A  
Patent No. 5773691  
GENERAL INFORMATION:  
APPLICANT: E. I. DU PONT DE NEMOURS AND  
APPLICANT: COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND  
METHODS FOR INCREASING  
TITLE OF INVENTION: INCREASING THE LYSINE  
TITLE OF INVENTION: AND THREONINE CONTENT  
TITLE OF INVENTION: OF THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS  
ADDRESSEE: AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD VERSION 2.0C  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,633A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: BARBARA C. SIEGELL  
REGISTRATION NUMBER: 30,684  
REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..28  
OTHER INFORMATION: /label= name  
OTHER INFORMATION: /note= "(SSP 7)4"

US-08-474-633A-70

Query Match 30.3%; Score 40; DB 1; Length 28;  
Best Local Similarity 25.0%; Pred. No. 38;  
Matches 5; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQKIDALE 20  
DB 4 KKKAMEEKKLAMEEKLKAME 23

RESULT 10  
US-08-823-771-70

APPLICANT: Henderson, Robert A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.470C9  
CURRENT APPLICATION NUMBER: US/09/834,759  
CURRENT FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 547  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 534  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-834-759-534

Query Match 31.1%; Score 41; DB 4; Length 21;  
Best Local Similarity 50.0%; Pred. No. 20;  
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 LKQKIASLKQKIDALEYE 22  
DB 4 LKKEIAMLKLEIATLKHQ 21

RESULT 8  
US-08-182-175A-3  
Sequence 3, Application US/08182175A  
Patent No. 5559223  
GENERAL INFORMATION:  
APPLICANT: Saverio Carl Falco  
APPLICANT: Sharon J. Keeler  
APPLICANT: Janet A. Rice  
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing B  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. du Pont de Nemours and Company  
STREET: 1007 Market Street  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/182,175A  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/743,006  
FILING DATE: 9 August 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Linda Axamethy Floyd  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: BB-1031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 992-4929  
TELEFAX: (302) 892-7949  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..28  
OTHER INFORMATION: /label= name  
OTHER INFORMATION: /note= "(SSP 7)4"

US-08-182-175A-3

Query Match 30.3%; Score 40; DB 1; Length 28;  
Best Local Similarity 25.0%; Pred. No. 38;  
Matches 5; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQKIDALE 20  
DB 4 KKKAMEEKKLAMEEKLKAME 23

RESULT 9  
US-08-474-633A-70  
Sequence 70, Application US/08474633A  
Patent No. 5773691  
GENERAL INFORMATION:  
APPLICANT: E. I. DU PONT DE NEMOURS AND  
APPLICANT: COMPANY  
TITLE OF INVENTION: CHIMERIC GENES AND  
METHODS FOR INCREASING  
TITLE OF INVENTION: INCREASING THE LYSINE  
TITLE OF INVENTION: AND THREONINE CONTENT  
TITLE OF INVENTION: OF THE SEEDS OF PLANTS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS  
ADDRESSEE: AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD VERSION 2.0C  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,633A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: BARBARA C. SIEGELL  
REGISTRATION NUMBER: 30,684  
REFERENCE/DOCKET NUMBER: BB-1037-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-992-4931  
TELEFAX: 302-773-0164  
TELEX: 835420  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..28  
OTHER INFORMATION: /label= name  
OTHER INFORMATION: /note= "(SSP 7)4"

US-08-474-633A-70

Query Match 30.3%; Score 40; DB 1; Length 28;  
Best Local Similarity 25.0%; Pred. No. 38;  
Matches 5; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQKIDALE 20  
DB 4 KKKAMEEKKLAMEEKLKAME 23

RESULT 10  
US-08-823-771-70

; Sequence 70, Application US/08823771  
; Patent No. 6459019  
; GENERAL INFORMATION:  
; APPLICANT: E. I. DU PONT DE NEMOURS AND  
; COMPANY  
; TITLE OF INVENTION: CHIMERIC GENES AND  
; METHODS FOR INCREASING  
; INCREASING THE LYSINE  
; AND THREONINE CONTENT  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: E. I. DU PONT DE NEMOURS  
; AND COMPANY  
; STREET: 1007 MARKET STREET  
; CITY: WILMINGTON  
; STATE: DELAWARE  
; COUNTRY: U.S.A.  
; ZIP: 19898  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MICROSOFT WORD VERSION 2.0C  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823,771  
; FILING DATE: 24-Mar-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/474,633  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BARBARA C. SIEGELL  
; REGISTRATION NUMBER: 30,684  
; REFERENCE/DOCKET NUMBER: BB-1037-C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302-992-4931  
; TELEFAX: 302-773-0164  
; TELEX: 835420  
; INFORMATION FOR SEQ ID NO: 70:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..28  
; OTHER INFORMATION: /label= name  
; /note= "(SSP 7)4"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 70:  
US-08-823-771-70  
Query Match 30.3%; Score 40; DB 4; Length 28;  
Best Local Similarity 25.0%; Pred. No. 38;  
Matches 5; Conservative 12; Mismatches 3; Indels 0; Gaps 0;  
Oy 1 KIAALKQKIASLKQEIADALE 20  
|:|:|:|:|:|:|:|:|:|:  
Db 4 KLKAMEEKLKAMEEKLKAME 23  
RESULT 11  
PCT-US92-06412-3  
; Sequence 3, Application PC/TUS9206412  
; GENERAL INFORMATION:  
; APPLICANT: Saverio Carl Falco  
; APPLICANT: Sharon J. Keeler  
; APPLICANT: Janet A. Rice  
; TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing P  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: E.I. du Pont de Nemours and Company

; STREET: 1007 Market Street  
; CITY: Wilmington  
; STATE: Delaware  
; COUNTRY: USA  
; ZIP: 19898  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: Macintosh System, 6.0  
; SOFTWARE: Microsoft Word, 4.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/06412  
; FILING DATE: 19920807  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/743,006  
; FILING DATE: 9 August 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Linda Axamethy Floyd  
; REGISTRATION NUMBER: 33,692  
; REFERENCE/DOCKET NUMBER: BB-1031  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (302) 992-4929  
; TELEFAX: (302) 892-7949  
; TELEX: 835420  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..28  
; OTHER INFORMATION: /label= name  
; OTHER INFORMATION: /note= "(SSP 7)4"  
PCT-US92-06412-3  
Query Match 30.3%; Score 40; DB 5; Length 28;  
Best Local Similarity 25.0%; Pred. No. 38;  
Matches 5; Conservative 12; Mismatches 3; Indels 0; Gaps 0;  
Oy 1 KIAALKQKIASLKQEIADALE 20  
|:|:|:|:|:|:|:|:|:|:  
Db 4 KLKAMEEKLKAMEEKLKAME 23  
RESULT 12  
US-08-737-629-8  
; Sequence 8, Application US/08737629  
; Patent No. 6190886  
; GENERAL INFORMATION:  
; APPLICANT: Hoppe, Hans-Jurgen  
; APPLICANT: Reid, Kenneth BM  
; TITLE OF INVENTION: Trimerising polypeptides, their manufacture  
; TITLE OF INVENTION: and use.  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye PC  
; STREET: 8th Floor, 1100 No. 6190886th Glebe Road  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: United States of America  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/737,629  
; FILING DATE: 10-JAN-1997  
; CLASSIFICATION: 435

Thu Apr 28 07:26:21 2005

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01104
; FILING DATE: 16-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9409768.0
; FILING DATE: 16-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-737-629-8

Query Match      29.5%; Score 39; DB 3; Length 25;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY  9 IASLKQEIADALE 20
DB  1 VASLRQQVEALQ 12

RESULT 13
US-08-446-922-9
; Sequence 9, Application US/08446922
; Patent No. 5716805
; GENERAL INFORMATION:
; APPLICANT: Spriggs, Melanie
; APPLICANT: Srinivasan, Subhashini
; TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,922
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 08/107,353
; FILING DATE: 08-13-93
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 1003-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-446-922-9

Query Match      29.5%; Score 39; DB 1; Length 27;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY  9 IASLKQEIADALE 20
DB  3 VASLRQQVEALQ 14

RESULT 14
US-08-484-624A-24
; Sequence 24, Application US/08484624A
; Patent No. 5962406
; GENERAL INFORMATION:
; APPLICANT: ARMITAGE, RICHARD
; APPLICANT: FANSLAW, WILLIAM
; APPLICANT: SPRIGGS, MELANIE
; APPLICANT: SRINIVASAN, SUBHASHINI
; APPLICANT: GIBSON, MARYLOU
; APPLICANT: MORRIS, ARVIA E.
; APPLICANT: MCGREW, JEFFERY
; TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,624A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/477,733
; FILING DATE: June 07, 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/969,703
; FILING DATE: October 23, 1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/805,723
; FILING DATE: December 5, 1991
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/783,707
; FILING DATE: October 25, 1991
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2802-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 2065870430
; TELEFAX: 2065870606
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
```

; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-484-624A-24

Query Match 29.5%; Score 39; DB 2; Length 27;  
Best Local Similarity 50.0%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 IASLKQEIADALE 20  
:||||:|::|:  
Db 3 VASLRQQVEALQ 14

## RESULT 15

US-08-477-733B-24  
; Sequence 24, Application US/08477733B  
; Patent No. 5981724

; GENERAL INFORMATION:  
; APPLICANT: ARMITAGE, RICHARD  
; APPLICANT: FANSLAW, WILLIAM  
; APPLICANT: SPRIGGS, MELANIE  
; APPLICANT: SRINIVASAN, SUBHASHINI  
; APPLICANT: GIBSON, MARYLOU  
; APPLICANT: MORRIS, ARVIA E.  
; APPLICANT: MCGREW, JEFFERY  
; TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMUNEX CORPORATION  
; STREET: 51 UNIVERSITY STREET  
; CITY: SEATTLE  
; STATE: WASHINGTON  
; COUNTRY: USA  
; ZIP: 98101

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,733B

; FILING DATE: June 07, 1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/249,189

; FILING DATE: May 24, 1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/969,703

; FILING DATE: October 23, 1992

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/805,723

; FILING DATE: December 5, 1991

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/783,707

; FILING DATE: October 25, 1991

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Perkins, Patricia A.

; REGISTRATION NUMBER: 34,693

; REFERENCE/DOCKET NUMBER: 2802-D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 2065870430

; TELEFAX: 2065870606

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 27 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-477-733B-24

Query Match 29.5%; Score 39; DB 2; Length 27;  
Best Local Similarity 50.0%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 IASLKQEIADALE 20  
:||||:|::|:  
Db 3 VASLRQQVEALQ 14

Search completed: April 27, 2005, 15:19:42  
Job time : 38.5 secs

**BEST AVAILABLE COPY**

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2005, 14:54:59 ; Search time 23:5 Seconds  
(without alignments)  
114.641 Million cell updates/sec

Title: US-10-088-417A-1

Perfect score: 132

Sequence: 1 KIAALKQKIASLKQEIADALEYENDALEQ 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 5694

Minimum DB seq length: 0

Maximum DB seq length: 28

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	25.0	22	2 JP0069	ribosomal protein
2	33	25.0	25	2 S27229	prothymosin alpha
3	31	23.5	25	2 JP0067	ribosomal protein
4	30	22.7	20	2 A61276	superoxide dismuta
5	30	22.7	27	2 A38123	probable maud prot
6	29	22.0	16	2 B58503	superoxide dismuta
7	29	22.0	21	2 F38837	T-cell receptor be
8	29	22.0	22	2 JP0066	ribosomal protein
9	28	21.2	20	2 A09233	light meromyosin -
10	28	21.2	23	2 P00858	Maud protein - Par
11	27	20.5	21	2 I40659	hypothetical prote
12	27	20.5	21	2 I40657	hypothetical prote
13	27	20.5	25	2 A44790	probable colonizat
14	27	20.5	25	2 JP0065	ribosomal protein
15	27	20.5	28	2 A61417	bdellin B-3 - medi
16	26.5	20.1	26	2 F69265	hypothetical prote
17	26.5	20.1	27	2 S17646	alcohol dehydrogen
18	26	19.7	23	2 S24279	hypothetical prote
19	26	19.7	24	2 A43431	phosphorylase kina
20	26	19.7	25	2 A23605	histone H1.1 - whe
21	26	19.7	26	2 P10027	M protein pepm19 -
22	26	19.7	27	2 I46492	myosin - rabbit (f
23	25.5	19.3	20	2 A53592	H+-exporting ATPas
24	25	18.9	14	2 A61032	troponin T, cardia
25	25	18.9	15	2 PA0025	protein OA300026 -
26	25	18.9	19	2 PNO457	nitrogenase (EC 1.
27	25	18.9	24	2 S47281	hypothetical prote
28	25	18.9	25	2 B44560	terephthalate 1,2-
29	25	18.9	28	2 PL0005	pepsin A (EC 3.4.2

photosystem I chai  
fructose-bisphosph  
superoxide dismuta  
15K protein A - ra  
15K protein B - ra  
15K protein C - ra  
cyclolytic enteroco  
DNA-binding protei  
hypothetical prote  
myosin heavy chain  
cell surface prote  
T-cell receptor be  
Ig H chain V-D-J r  
tubulin alpha-chai  
mastoparan M - hor  
mastoparan - Yello  
histamine-releasin

## ALIGNMENTS

### RESULT 1

JP0069

ribosomal protein L30 - Thermomonospora mesophila (fragment)

C:Species: Thermomonospora mesophila

C>Date: 10-Mar-1994 #sequence\_revision 28-Oct-1994 #text\_change 28-Oct-1994

C:Accession: JP0069

R:Ochi, K.

submitted to JIPID, February 1994

A:Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal p

A:Reference number: JP0042

A:Accession: JP0069

A:Molecule type: protein

A:Residues: 1-22 <OCH>

C:Keywords: protein biosynthesis; ribosome

Query Match 25.0%; Score 33; DB 2; Length 22;

Best Local Similarity 36.8%; Pred. No. 5e+02;

Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KIAALKQKIASLKQEIADAL 19

Db 4 KITQLRSKIGKQKQDXDL 22

### RESULT 2

S27229

prothymosin alpha homolog - Escherichia coli (fragments)

C:Species: Escherichia coli

C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 10-Nov-1995

C:Accession: S27229

R:Vartapetian, A.; Chichkova, N.; Lyakhov, I.; Makarova, T.; Evstafieva, A.; Bogdanov, A.

FEBS Lett. 313, 95-97, 1992

A:Title: Segments of Escherichia coli genome similar to the exons of human prothymosin a

A:Reference number: S27229; MUID:93050251; PMID:1426289

A:Accession: S27229

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-6;7-25 <VAR>

A>Note: the authors did not translate the codon for residues 17

Query Match 25.0%; Score 33; DB 2; Length 25;

Best Local Similarity 42.1%; Pred. No. 5.8e+02;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 10 ASLKQEIADALEYENDALEQ 28

Db 4 ASVNSENGELEDNDDEE 22

### RESULT 3

JP0067

ribosomal protein L30 - Thermomonospora chromogena (fragment)

Thu Apr 28 07:26:22 2005

C;Species: Thermomonospora Chromogena  
C;Date: 10-Mar-1994 #sequence\_revision 28-Oct-1994 #text\_change 28-Oct-1994  
C;Accession: JP0067  
R;Ochi, K.  
submitted to JPIID, February 1994  
A;Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal protein  
A;Reference number: JP0042  
A;Accession: JP0067  
A;Molecule type: protein  
A;Residues: 1-25 <OCH>  
C;Keywords: protein biosynthesis; ribosome

Query Match 23.5%; Score 31; DB 2; Length 25;  
Best Local Similarity 31.6%; Pred. No. 1.e+03; Indels 0; Gaps 0;  
Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQIDAL 19  
DB 4 KITQVRKIGKGKQNXDSL 22

RESULT 4  
A61276  
superoxide dismutase (EC 1.15.1.1) (Fe/Mn) homolog - chicken (fragment)  
N;Alternate names: osteoclast membrane glycoprotein  
C;Species: Gallus gallus (chicken)  
C;Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 05-Mar-1999  
C;Accession: A61276  
J;Ouralier, M.J.; Li, L.; Osadoby, P.  
J;Cell. Biochem. 46, 219-233, 1991  
A;Title: Purification and characterization of an osteoclast membrane glycoprotein with h  
A;Reference number: A61276; MUID:92129474; PMID:1723067  
A;Accession: A61276  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-20 <OUR>  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Keywords: metalloprotein; oxidoreductase

Query Match 22.7%; Score 30; DB 2; Length 20;  
Best Local Similarity 53.8%; Pred. No. 1.e+03; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 15 EIDALEYENDALE 27  
DB 3 ELPDLPAYDALE 15

RESULT 5  
A38123  
probable maud protein - Methylobacterium extorquens (strain AM1) (fragment)  
C;Species: Methylobacterium extorquens  
C;Date: 24-Jul-1992 #sequence\_revision 24-Jul-1992 #text\_change 09-Jul-2004  
C;Accession: A38123  
R;Chistoserdov, A.Y.; Lidstrom, M.E.  
J;Bacteriol. 173, 5909-5913, 1991  
A;Title: The small-subunit polypeptide of methylamine dehydrogenase from Methylobacteriu  
A;Reference number: A38123; MUID:91358386; PMID:1855555  
A;Accession: A38123  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-27 <CHI>  
A;Cross-references: UNIPROT:Q49126; GB:M57963

Query Match 22.7%; Score 30; DB 2; Length 27;  
Best Local Similarity 29.2%; Pred. No. 1.5e+03; Indels 0; Gaps 0;  
Matches 7; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 2 IAALKQKIASLKQIDALEYENDA 25  
DB 1 LEADKSGFASIQFMTSRKSHDA 24

C;Species: Thermomonospora Chromogena  
C;Date: 10-Mar-1994 #sequence\_revision 28-Oct-1994 #text\_change 28-Oct-1994  
C;Accession: JP0067  
R;Ochi, K.  
submitted to JPIID, February 1994  
A;Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal protein  
A;Reference number: JP0042  
A;Accession: JP0067  
A;Molecule type: protein  
A;Residues: 1-25 <OCH>  
C;Keywords: protein biosynthesis; ribosome

Query Match 23.5%; Score 31; DB 2; Length 25;  
Best Local Similarity 31.6%; Pred. No. 1.e+03; Indels 0; Gaps 0;  
Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIAALKQKIASLKQIDAL 19  
DB 4 KITQVRKIGKGKQNXDSL 22

RESULT 4  
A61276  
superoxide dismutase (EC 1.15.1.1) (Fe/Mn) homolog - chicken (fragment)  
N;Alternate names: osteoclast membrane glycoprotein  
C;Species: Gallus gallus (chicken)  
C;Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 05-Mar-1999  
C;Accession: A61276  
J;Ouralier, M.J.; Li, L.; Osadoby, P.  
J;Cell. Biochem. 46, 219-233, 1991  
A;Title: Purification and characterization of an osteoclast membrane glycoprotein with h  
A;Reference number: A61276; MUID:92129474; PMID:1723067  
A;Accession: A61276  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-20 <OUR>  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Keywords: metalloprotein; oxidoreductase

Query Match 22.7%; Score 30; DB 2; Length 20;  
Best Local Similarity 53.8%; Pred. No. 1.e+03; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 15 EIDALEYENDALE 27  
DB 3 ELPDLPAYDALE 15

RESULT 5  
A38123  
probable maud protein - Methylobacterium extorquens (strain AM1) (fragment)  
C;Species: Methylobacterium extorquens  
C;Date: 24-Jul-1992 #sequence\_revision 24-Jul-1992 #text\_change 09-Jul-2004  
C;Accession: A38123  
R;Chistoserdov, A.Y.; Lidstrom, M.E.  
J;Bacteriol. 173, 5909-5913, 1991  
A;Title: The small-subunit polypeptide of methylamine dehydrogenase from Methylobacteriu  
A;Reference number: A38123; MUID:91358386; PMID:1855555  
A;Accession: A38123  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-27 <CHI>  
A;Cross-references: UNIPROT:Q49126; GB:M57963

Query Match 22.7%; Score 30; DB 2; Length 27;  
Best Local Similarity 29.2%; Pred. No. 1.5e+03; Indels 0; Gaps 0;  
Matches 7; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 2 IAALKQKIASLKQIDALEYENDA 25  
DB 1 LEADKSGFASIQFMTSRKSHDA 24

## RESULT 6

E58503  
superoxide dismutase (EC 1.15.1.1) - unidentified bacterium (fragment)  
N;Alternate names: 21.3K bladder and kidney stone protein  
C;Species: unidentified bacterium  
C;Date: 07-Feb-1997 #sequence\_revision 07-Feb-1997 #text\_change 09-Jul-2004  
C;Accession: E58503  
R;Binette, J.P.; Binette, M.B.  
submitted to the Protein Sequence Database, October 1996  
A;Description: The proteins of kidney and gallbladder stones.  
A;Reference number: A58501  
A;Accession: E58503  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-16 <BIN>  
A;Cross-references: UNIPROT:Q7M137  
A;Experimental source: human bladder and kidney stones  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Keywords: metalloprotein; oxidoreductase

Query Match 22.0%; Score 29; DB 2; Length 16;  
Best Local Similarity 40.0%; Pred. No. 1.2e+03; Indels 0; Gaps 0;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 12 LKQIDALEYENDAL 26

DB 1 MEHTLPPPYEMDAL 15

## RESULT 7

F38837  
T-cell receptor beta chain precursor V region (23.1) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 23-Jul-1999  
C;Accession: F38837  
R;Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.; Panayi, G.;  
Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991  
A;Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheumat  
A;Reference number: A41299; MUID:92020887; PMID:1656449  
A;Accession: F38837  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-21 <UEM>  
A;Cross-references: GB:S57606; NID:G236346; PIDN:AA19970.1; PID:G236347  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: T-cell receptor

Query Match 22.0%; Score 29; DB 2; Length 21;  
Best Local Similarity 40.0%; Pred. No. 1.5e+03; Indels 0; Gaps 0;  
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 12 LKQIDALEYENDAL 26

DB 1 LSLNVNALLEDDAL 15

## RESULT 8

JP0066  
ribosomal protein L30 - Nocardia asteroides (fragment)  
C;Species: Nocardia asteroides  
C;Date: 10-Mar-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004  
C;Accession: JP0066  
R;Ochi, K.  
submitted to JPIID, February 1994  
A;Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal p  
A;Reference number: JP0042  
A;Accession: JP0066  
A;Molecule type: protein  
A;Residues: 1-22 <OCH>  
A;Cross-references: UNIPROT:Q7M028  
C;Keywords: protein biosynthesis; ribosome



```
Query Match      22.0%; Score 29; DB 2; Length 22;
Best Local Similarity 31.6%; Pred. No. 1.6e+03;
Matches 6; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 8 KIASLKQBIIDALEYENDAL 26
   : : | | | : : | |
DB 4 KVTQIKSTIGAKANQKDSL 22

RESULT 9
A20923
light meromyosin - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C:Accession: A20923
R:Nyitrai, L.; Mocz, G.; Szilagyi, L.; Balint, M.; Lu, R.C.; Wong, A.; Gergely, J.
J. Biol. Chem. 258, 13213-13220, 1983
A:Title: The proteolytic substructure of light meromyosin.
A:Reference number: A92403; MUID:84032553; PMID:6355107
A:Accession: A20923
A:Molecule type: protein
A:Residues: 1-18 <NYI>
A:Cross-references: UNIPROT:Q8JG72
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: ATP; muscle

Query Match      21.2%; Score 28; DB 2; Length 18;
Best Local Similarity 35.7%; Pred. No. 1.7e+03;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 ALKQKIASLKQBIID 17
   | | | | | | | |
DB 4 AFTQKIELEKQKLE 17

RESULT 10
PH0858
Maud protein - Paracoccus denitrificans (fragment)
C:Species: Paracoccus denitrificans
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 09-Jul-2004
C:Accession: PH0858
R:Chistoserdov, A.Y.; Boyd, J.; Mathews, F.S.; Lidstrom, M.E.
Biochem. Biophys. Res. Commun. 184, 1181-1189, 1992
A:Title: The genetic organization of the mau gene cluster of the facultative autotroph B
A:Reference number: PH0858; MUID:92272706; PMID:1590782
A:Accession: PH0858
A:Molecule type: DNA
A:Residues: 1-23 <CHI>
A:Cross-references: UNIPROT:P29895; GB:M90098; NID:g150580; PIDN:AAA25577.1; PID:g150581
C:Genetics:
A:Gene: maud

Query Match      21.2%; Score 28; DB 2; Length 23;
Best Local Similarity 46.2%; Pred. No. 2.2e+03;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 AALKQKIASLKQBE 15
   | : | : | | | :
DB 9 ASLQOQYMASRKQK 21

RESULT 11
I40659
hypochemical protein - Clostridium cochlearium (fragment)
C:Species: Clostridium cochlearium
C>Date: 12-Aug-1996 #sequence_revision 18-Oct-1996 #text_change 09-Jul-2004
C:Accession: I40659; S47461
R:Zelder, O.; Beatrice, B.; Leutbecher, U.; Buckel, W.
Eur. J. Biochem. 226, 577-585, 1994
A:Title: Characterization of the coenzyme-B12-dependent glutamate mutase from Clostridium
A:Reference number: I40659; MUID:95094816; PMID:7880251
A:Accession: I40659

A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-21 <RES>
A:Cross-references: UNIPROT:Q46011; EMBL:X80997; NID:g530005; PIDN:CAA56920.1; PID:g530005

Query Match      20.5%; Score 27; DB 2; Length 21;
Best Local Similarity 41.7%; Pred. No. 2.7e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IAALKQKIASLK 13
   : | | | | : |
DB 7 VKALKSKLEVK 18

RESULT 12
I40657
hypochemical protein 21 - Clostridium cochlearium (fragment)
C:Species: Clostridium cochlearium
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: I40657; S41331
R:Zelder, O.; Beatrice, B.; Buckel, W.
FEMS Microbiol. Lett. 118, 15-21, 1994
A:Title: Cloning, sequencing and expression in Escherichia coli of the gene encoding con
A:Reference number: I40657; MUID:94283856; PMID:8013871
A:Accession: I40657
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-21 <RES>
A:Cross-references: UNIPROT:Q46010; EMBL:X75890; NID:g441487; PIDN:CAA53483.1; PID:g441487

Query Match      20.5%; Score 27; DB 2; Length 21;
Best Local Similarity 41.7%; Pred. No. 2.7e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IAALKQKIASLK 13
   : | | | | : |
DB 7 VKALKSKLEVK 18

RESULT 13
A44790
probable colonization factor O166 - Escherichia coli (fragment)
C:Species: Escherichia coli
C>Date: 24-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A44790
R:Sommerfeld, H.; Grewal, H.M.; Svennerholm, A.M.; Gastra, W.; Flood, P.R.; Viboud, C.,
Infect. Immun. 60, 3799-3806, 1992
A:Title: Genetic relationship of putative colonization factor O166 to colonization facto
A:Reference number: A44790; MUID:92363580; PMID:1354200
A:Contents: E7476A
A:Accession: A44790
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <SOM>
A:Cross-references: UNIPROT:Q9R5J4
A:Note: sequence extracted from NCBI backbone (NCBI:111010)
C:Superfamily: CPA1 fibrial protein

Query Match      20.5%; Score 27; DB 2; Length 25;
Best Local Similarity 41.2%; Pred. No. 3.2e+03;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 10 ASLKQEIIDALEYENDAL 26
   | : | : | | | :
DB 9 ASVDPITDILQNGSAL 25

RESULT 14
JP0065
ribosomal protein L30 - Streptomyces rochei (fragment)
C:Species: Streptomyces rochei
C>Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 21-Aug-1998
C:Accession: JP0065
```

R;Ochi, K.

submitted to JIPID, February 1994

**A;Description:** phylogenetic diversity in the genus *Bacillus* and comparative ribosomal pr

A;Reference number: JP0042

A;Accession: JP0065

A;Molecule type: protein  
A:Residues: 1-27 <OCH>

A/Residues: 1-27 <OCH>  
C: Superfamily: Escherichia coli ribosomal protein L30

**C;Keywords:** protein biosynthesis; ribosome

Query Match 20.5%; score 27; DB 2; Length 27;

Best Local Similarity	36.8%;	Pred. No. 3.5e+03;
-----------------------	--------	--------------------

Matches	7;	Conservative	2;	Mismatches	10;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	-----	--------	----	------	----

Qy 1 KIAALKQKIASLKQEIDAL 19

4 7 T T A I W S E V I C C A V N V V D T I . 22

## RESULT 15

A61417

bdellin B-3 - medicinal leech (fragment)

C;Species: Hirudo medicinalis (medicinal leech)  
C;2-5-1994 #accessions 08-Sep-1994 #text change 07-May-1999

```
C;date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1993
C;Accession: A61417
```

C;Accession: A6141/  
P.Kreici, K.: Fritz, H.

K;NIEJC, A.; FLECH, H.  
FEBS Lett. 64, 152-155, 1976

A,Title: Structural homology of a trypsin-plasmin inhibitor from leeches (bdellin B-3) w

A;Reference number: A61417; MUID:76188026; PMID:131707

A;Accession: A61417

**A;Status: preliminary**

A;Molecule type: protein  
A.Residues: 1-28 (KRF)

A;REBIQUEB: I-28 <NRE>

Query Match	20.58; Score 27; DB 2; Length 28;

**Best Local Similarity** 40.0%; **Pred. No.** 3.6e+03;

Matches	4;	Conservative	3;	Mismatches	3;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

17 DALEYENDAL 26 Ov

QY	DATE	TIME	NO.
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30
31	31	31	31
32	32	32	32
33	33	33	33
34	34	34	34
35	35	35	35
36	36	36	36
37	37	37	37
38	38	38	38
39	39	39	39
40	40	40	40
41	41	41	41
42	42	42	42
43	43	43	43
44	44	44	44
45	45	45	45
46	46	46	46
47	47	47	47
48	48	48	48
49	49	49	49
50	50	50	50
51	51	51	51
52	52	52	52
53	53	53	53
54	54	54	54
55	55	55	55
56	56	56	56
57	57	57	57
58	58	58	58
59	59	59	59
60	60	60	60
61	61	61	61
62	62	62	62
63	63	63	63
64	64	64	64
65	65	65	65
66	66	66	66
67	67	67	67
68	68	68	68
69	69	69	69
70	70	70	70
71	71	71	71
72	72	72	72
73	73	73	73
74	74	74	74
75	75	75	75
76	76	76	76
77	77	77	77
78	78	78	78
79	79	79	79
80	80	80	80
81	81	81	81
82	82	82	82
83	83	83	83
84	84	84	84
85	85	85	85
86	86	86	86
87	87	87	87
88	88	88	88
89	89	89	89
90	90	90	90
91	91	91	91
92	92	92	92
93	93	93	93
94	94	94	94
95	95	95	95
96	96	96	96
97	97	97	97
98	98	98	98
99	99	99	99
100	100	100	100

Search completed: April 27, 2005, 15:07:07

Job time : 29.5 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	38	28.8	20	1	SODP_PASPI	P81527	pasteurella
2	38	28.8	23	2	90UCS1	Q9UCe1	homo sapien
3	35	26.5	26	2	Q7RR69	plasmodium	Q7rr69
4	34	25.8	22	2	O44710	drosophila	O44710
5	32	24.2	16	2	Q95Y34	caenorhabdi	Q95y34
6	32	24.2	28	2	Q8ZYI3	pyrobaculum	Q8zyi3
7	31	23.5	15	1	SODM_ENTAE	enterobacte	P22799
8	31	23.5	18	2	Q9QZK4	mus spretus	Q9qzx4
9	31	23.5	18	2	Q9QZK5	mus musculus	Q9qzx5
10	31	23.5	19	2	Q9UC80	homo sapien	Q9uc80
11	31	23.5	27	2	Q9NUT6	leishmania	Q9njt6
12	30	22.7	28	2	O65311	arabidopsis	O65311
13	29	22.0	16	2	Q7M137	unidentifie	Q7m137
14	29	22.0	19	1	OXLA_OPHHA	ophiophagus	P81383
15	29	22.0	22	2	Q7M0Z8	nocardia as	Q7m0z8
16	29	22.0	25	2	Q7FCW4	plasmodium	Q7pcw4
17	29	22.0	25	2	Q7RLH1	plasmodium	Q7rlh1
18	29	22.0	26	2	Q9CS32	mus musculus	Q9cs32
19	29	22.0	27	2	Q25867	plasmodium	Q25867
20	29	22.0	27	2	O661K4	borrelia ga	Q661k4
21	28.5	21.6	24	2	Q9RQ27	clostridium	Q9rq27
22	28	21.2	15	2	Q6TAR1	homo sapien	Q6tar1
23	28	21.2	21	2	Q638L3	rattus sp.	Q638l3
24	28	21.2	22	2	Q91QR9	cauliflower	Q91qr9
25	28	21.2	23	2	Q9UWL6	methanosarc	Q9uwl6
26	28	21.2	24	2	O9GNE5	caenorhabdi	Q9gne5
27	28	21.2	26	2	Q9UWF9	methanosarc	Q9uwf9
28	28	21.2	26	2	Q71QT7	flavescence	Q71qt7
29	28	21.2	26	2	Q99JC5	rattus sp.	Q99jc5
30	28	21.2	28	2	Q9RAT7	lactococcus	Q9rat7
31	27	20.5	15	2	Q6LDF5	trichostrotr	Q6ldf5

AC	044710.	
DT	01-JUN-1998 (TReMBLrel. 06, Created)	
DT	01-JUN-1998 (TReMBLrel. 06, Last sequence update)	
DT	01-JUN-2001 (TReMBLrel. 17, Last annotation update)	
DE	Tropomyosin 1 isoform e (Fragment).	
DE	Name=Tml;	
OS	Drosophila pseudoobscura (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7237;	
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=Goldendale 3;	
RP	MEDLINE=99442389; PubMed=10511563;	
RX	Hamlin M.T., Aquadro C.F.;	
RA	"DNA sequence variation and the recombinational landscape in	
RA	Drosophila pseudoobscura. A study of the second chromosome.";	
RT	Genetics 153:859-869 (1999).	
RL	EMBL; AF039273; AAB96670.1; -	
DR	EMBL; AF039273; AAB96670.1; JOINED.	
DR	FlyBase; FBgn0025413; Dpse\Tml.	
DR	InterPro; IPR000533; Tropomyosin.	
DR	Pfam; PF00261; tropomyosin; 1.	
FT	NON_TER 1	
FT	NON_TER 22	
FT	SEQUENCE 22 AA; 2611 MW; F86B844608F4175C CRC64;	
SQ		
Query Match	25.88; Score 34; DB 2; Length 22;	
Best Local Similarity	35.3%; Pred.No. 3.3e+03;	
Matches	6; Conservative 7; Mismatches 4; Indels 0; Gaps	
QY	6 KQIASLKQEIADLEYE 22	
DB	==== : : ====	
DB	2 EKQVKRLQKEVDLEDE 18	
RESULT 5		
Q95Y34	PRELIMINARY; PRT; 16 AA.	
ID	Q95Y34	
AC	095Y34;	
DT	01-DEC-2001 (TReMBLrel. 19, Created)	
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)	
DT	01-OCT-2003 (TReMBLrel. 25, Last annotation update)	
DE	Hypothetical protein Y108G3AL.6.	
DE	Name=Y108G3AL.6; ORFNames=Y108G3AL.6;	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	
OC	Rhabditidae; Peloderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=Bristol N2;	
RP	MEDLINE=99069613; PubMed=9851916;	
RX	WormBase Consortium;	
RG	"Genome sequence of the nematode C. elegans: a platform for	
RT	investigating biology. The C. elegans Sequencing Consortium.";	
RL	Science 282:2012-2018 (1998).	
RL	[2]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=Bristol N2;	
RP	Courtney L.;	
RA	"The sequence of C. elegans cosmid Y108G3AL.";	
RT	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
RL	[3]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=Bristol N2;	
RP	Waterston R.H.;	
RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
RL	[4]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=Bristol N2;	
RP	Waterston R.;	
RA		

RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA WormBase Consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AC024744; AALJ3326.1; --  
 DR WormBase; WBGene00022437; Y108G3AL.6.  
 DR WormPep; Y108G3AL.6; CE29605.  
 KW Hypothetical protein.  
 SQ SEQUENCE 16 AA; 1904 MW; 4932099B5AA3161E5 CRC64;  
  
 Query Match 24.2%; Score 32; DB 2; Length 16;  
 Best Local Similarity 35.7%; Pred. No. 4.2e+03;  
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
  
 Qy 15 EIDALEYENDALEQ 28  
 Db 2 EVDDSDYEDSDIEE 15  
  
 RESULT 6  
 Q82Y3 PRELIMINARY; PRT; 28 AA.  
 ID Q82Y3  
 AC Q82Y3;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein PAE0551.  
 GN OrderedLocusNames=PAE0551;  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2;  
 RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636499;

RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 RT aerophilum".  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
 DR EMBL; AE009773; AAL62858.1; --  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 28 AA; 3001 MW; 869F81422C5A14D CRC64;  
  
 Query Match 24.2%; Score 32; DB 2; Length 28;  
 Best Local Similarity 33.3%; Pred. No. 7.2e+03;  
 Matches 6; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
  
 Qy 2 IAAKQKIASLKQKIDAL 19  
 Db 3 VAEIKASVAELKVAVGSL 20  
  
 RESULT 7  
 SODM\_ENTAE STANDARD; PRT; 15 AA.  
 ID SODM\_ENTAE  
 AC P22759;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Superoxide dismutase [Fe] (EC 1.15.1.1) (Fragment).  
 GN Name=sodA;  
 OS Enterobacter aerogenes (Aerobacter aerogenes).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Enterobacter.  
 OX NCBI\_TaxID=548;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=91248479; PubMed=1368658;  
 RA Kim S.W., Lee S.O., Lee T.H.;  
 RT "Purification and characterization of superoxide dismutase from  
 RT Aerobacter aerogenes".  
 RL Agric. Biol. Chem. 55:101-108(1991).  
 CC -!- FUNCTION: Destroys radicals which are normally produced within the  
 CC cells and which are toxic to biological systems.  
 CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -!- COFACTOR: Binds 1 iron ion per subunit (By similarity).  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase  
 CC family.  
 DR PIR; P00615; P00615.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; Sod\_Fe\_N; 1.  
 DR PROSITE; PS00088; SOD\_MN; PARTIAL.  
 KW Direct protein sequencing; Iron; Metal-binding; Oxidoreductase.  
 FT NON TER 15  
 SQ SEQUENCE 15 AA; 1756 MW; 352F3D949202E642 CRC64;  
  
 Query Match 23.5%; Score 31; DB 1; Length 15;  
 Best Local Similarity 53.8%; Pred. No. 5.2e+03;  
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
 Qy 15 EIDALEYENDALE 27  
 Db 3 ELPQLPYAYDALE 15  
  
 RESULT 8  
 Q9QZX4 PRELIMINARY; PRT; 18 AA.  
 ID Q9QZX4  
 AC Q9QZX4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Desmoplakin (Fragment).  
 GN Name=Dep;  
 OS Mus spretus (Western wild mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10096;
RN [1]
RP SEQUENCE FROM N.A.
RA Green K.J., Guy S.G., Ceerhalmi-Friedman P., McLean W.H.I.,
RA Christiano A.M., Wagner R.M.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF148515; AAF02528.1; -.
DR MGD; MGI:109611; Dep.
FT NON_TER 1
FT NON_TER 18
SQ SEQUENCE 18 AA; 2289 MW; 3B41DF23C1E2F960 CRC64;

Query Match 23.5%; Score 31; DB 2; Length 18;
Best Local Similarity 41.2%; Pred. No. 6.2e+03;
Matches 7; Conservative 6; Mismatches 2; Indels 2; Gaps 1;

QY 11 SLKQEIIDALEYENDALE 27
Db 4 NLRQEIE--KFQKQALE 18

RESULT 9
Q9QZX5 PRELIMINARY; PRT; 18 AA.
AC Q9QZX5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Desmoplakin (Fragment).
GN Name-Dsp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C75BL6/J;
RA Green K.J., Guy S.G., Ceerhalmi-Friedman P., McLean W.H.I.,
RA Christiano A.M., Wagner R.M.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF148514; AAF02527.1; -.
DR MGD; MGI:109611; Dep.
DR GO; GO:0016323; C:bacolateral plasma membrane; IDA.
DR GO; GO:0030057; C:desmosome; IDA.
DR GO; GO:0005739; C:mitochondrion; IDA.
FT NON_TER 1
FT NON_TER 18
SQ SEQUENCE 18 AA; 2289 MW; 3B41DF23C1E2F960 CRC64;

Query Match 23.5%; Score 31; DB 2; Length 18;
Best Local Similarity 41.2%; Pred. No. 6.2e+03;
Matches 7; Conservative 6; Mismatches 2; Indels 2; Gaps 1;

QY 11 SLKQEIIDALEYENDALE 27
Db 4 NLRQEIE--KFQKQALE 18

RESULT 10
Q9UC80 PRELIMINARY; PRT; 19 AA.
AC Q9UC80;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE 21.5 kDa stone matrix protein (SC 1.15.1.1) (Superoxide dismutase [Mn/Fe]) (Fragment).
DE [Mn/Fe] (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metazoa; Chordata; Craniata; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE.
RA Binette J.P., Binette M.B.;
RT "Sequencing of proteins extracted from stones.";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (by similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR HSSP; P09223; lDT0.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF00081; Sod_Fe_N; 1.
KW Oxidoreductase.
SQ SEQUENCE 19 AA; 2116 MW; A0DDDAE848EE7894 CRC64;

Query Match 23.5%; Score 31; DB 2; Length 19;
Best Local Similarity 46.2%; Pred. No. 6.5e+03;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 15 EIDALEYENDALE 27
Db 3 ELPLPYAHDAIQ 15

RESULT 11
Q9NJ76 PRELIMINARY; PRT; 27 AA.
AC Q9NJ76;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major paraflagellar rod protein (fragment).
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RA Ch'ang L.Y., Lee T.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF110672; AAF66093.1; -.
DR GO; GO:0019861; C:flagellum; IEA.
KW Flagellum.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3005 MW; 0C5ADB2B6B425ED9 CRC64;

Query Match 23.5%; Score 31; DB 2; Length 27;
Best Local Similarity 53.8%; Pred. No. 9.1e+03;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 IAAKQKTASLKQ 14
Db 1 IVALKKTLLNLKQ 13

RESULT 12
O65311 PRELIMINARY; PRT; 28 AA.
AC O65311;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Beta-9 tubulin (Fragment).
GN Name-TUB9;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

```

```

RN SEQUENCE FROM N.A.
RP Sanders P.M., Bui A.O., Weterings K., McIntire K.N., Hau Y.C.,
RA Lee P.Y., Truong M.T., Beals T.B., Goldberg R.B.;
RT "Anther Development Defects in Arabidopsis thaliana Male-Sterile
RT Mutants.";
RL Sex. Plant Reprod. 11:297-322(1999).
DR EMBL; AF060248; AAC97107.1; -.
FT NON_TER 1
SQ SEQUENCE 28 AA; 3388 MW; 8060B4E537B5E70E CRC64;

Query Match 22.7%; Score 30; DB 2; Length 28;
Best Local Similarity 37.5%; Pred. No. 1.2e+04;
Matches 9; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

Qy 9 IASLKQEIADL---EYENDALEQ 28
Db 3 VAEYQYQDQATVGEYEEDEEE 26

RESULT 13
Q7M137
ID Q7M137 PRELIMINARY; PRT; 16 AA.
AC Q7M137;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Superoxide dismutase (EC 1.15.1.1) (Fragment).
OS unidentified bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=2338;
RN SEQUENCE.
RA Binette J.P., Binette M.B.;
RL Submitted (OCT-1996) to the PIR data bank.
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR PIR; E58503; E58503.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF00081; Sod_Fe_N; 1.
KW Oxidoreductase.
FT NON_TER 16
SQ SEQUENCE 16 AA; 1828 MW; 12DE78949AC43609 CRC64;

Query Match 22.0%; Score 29; DB 2; Length 16;
Best Local Similarity 40.0%; Pred. No. 9.5e+03;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 12 LKQEIDALEYENDAL 26
Db 1 MEHTLPPLPYENDAL 15

RESULT 14
OXLA OPHHA STANDARD; PRT; 19 AA.
AC P81383;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE L-amino-acid oxidase (EC 1.4.3.2) (LAO) (LAO) (Fragment).
OS Ophiophagus hannah (King cobra) (Naja hannah).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Ophiophagus.
OX NCBI_TaxID=8665;

```

```

RN SEQUENCE.
RP TISSUE=Venom;
EX MEDLINE=94361525; PubMed=8080286; DOI=10.1006/abbi.1994.1401;
RA Ponnudurai G., Chung M.C.M., Tan N.-H.;
RT "Purification and properties of the L-amino acid oxidase from Malayan
RT pit viper (Calloselasma rhodostoma) venom.";
RL Arch. Biochem. Biophys. 313:373-378(1994).
RN SEQUENCE OF 1-15.
RP TISSUE=Venom;
EX MEDLINE=97449790; PubMed=9304806; DOI=10.1016/S1357-2725(97)00024-1;
RA Ahn M.Y., Lee B.M., Kim Y.S.;
RT "Characterization and cytotoxicity of L-amino acid oxidase from the
RT venom of king cobra (Ophiophagus hannah).";
RL Int. J. Biochem. Cell Biol. 29:911-919(1997).
CC -!- FUNCTION: Has cytotoxic activity (By similarity).
CC -!- CATALYTIC ACTIVITY: An L-amino acid + H(2)O + O(2) = a 2-oxo acid
CC + NH(3) + H(2)O(2).
CC -!- COFACTOR: FAD.
CC -!- SUBUNIT: Homodimer (Probable).
CC -!- PTM: Glycosylated.
CC -!- SIMILARITY: Belongs to the flavin monoamine oxidase family.
CC Strong, to mammalian FGL1.
KW Direct protein sequencing; FAD; Flavoprotein; Glycoprotein;
KW Oxidoreductase; Toxin.
FT CONFLICT 1 H -> S (in Ref. 2).
FT NON_TER 19
SQ SEQUENCE 19 AA; 2298 MW; DD911A5B414F1427 CRC64;

Query Match 22.0%; Score 29; DB 1; Length 19;
Best Local Similarity 33.3%; Pred. No. 1.1e+04;
Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 9 IASLKQEIDALEYEN 23
Db 2 VINLEESQPEYEN 16

RESULT 15
Q7M028
ID Q7M028 PRELIMINARY; PRT; 22 AA.
AC Q7M028;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Ribosomal protein L30 (Fragment).
OS Nocardia asteroides.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=1824;
RN SEQUENCE.
RA Ochi K.;
RL Submitted (FEB-1994) to the PIR data bank.
DR PIR; JP0066; JP0066.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2330 MW; 5C7E71FB6FD577CA CRC64;

Query Match 22.0%; Score 29; DB 2; Length 22;
Best Local Similarity 31.6%; Pred. No. 1.3e+04;
Matches 6; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 8 KIASLKQEIDALEYENDAL 26
Db 4 KVTQIKSTIGAKNQKDSL 22

```

Search completed: April 27, 2005, 15:06:16  
Job time : 122 secs

**BEST AVAILABLE COPY**

**THIS PAGE LEFT BLANK**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2005, 14:45:41 ; Search time 122 Seconds  
(without alignments)  
88.765 Million cell updates/sec

Title: US-10-088-417A-4

Perfect score: 137  
Sequence: 1 KIRALKWNAHLKQEIQAIEQIALEQ 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 794064

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137	100.0	28	4	Aab74346 Peptide S
2	137	100.0	28	4	Aab74347 Peptide S
3	124	90.5	28	4	Aab74349 Peptide S
4	123	89.8	28	4	Aab74348 Peptide S
5	123	89.8	28	4	Aab74356 Peptide u
6	123	89.8	28	4	Aab74352 Peptide S
7	122	89.1	28	4	Aab74345 Peptide S
8	105	76.6	28	8	ADM41438 Self-asse
9	100	73.0	27	8	ADM41437 Self-asse
10	81	59.1	28	7	ABR84737 DE novo d
11	75	54.7	28	4	Aab74343 Peptide S
12	75	54.7	28	4	Aab74351 Peptide S
13	75	54.7	28	8	ADM41436 Self-asse
14	74	54.0	28	4	Aab74341 Peptide S
15	74	54.0	28	4	Aab74342 Peptide S
16	70	51.1	28	4	Aab74355 Peptide u
17	70	51.1	28	4	Aab74350 Peptide S
18	68	49.6	25	3	Aab08382 Peptide u
19	59	43.1	28	7	ABR84735 DE novo d
20	55	40.1	18	4	Aab74340 Peptide C
21	55	40.1	25	3	Aab08381 Peptide u
22	54	39.4	28	7	ABR84736 DE novo d
23	51	37.2	24	3	ABD21683 Coiled co
24	51	37.2	24	7	ADD93917 Coiled co
25	50	36.5	23	3	Aab08388 Peptide u

26	50	36.5	23	6	ADA00699	Ada00699 Protein m
27	47	34.3	14	2	AAW71411	Aaw71411 Peptide f
28	47	34.3	14	2	AAW80528	Aaw80528 Beta-shee
29	47	34.3	15	2	AAW71417	Aaw71417 Peptide f
30	47	34.3	15	2	AAW80545	Aaw80545 Peptide h
31	47	34.3	17	2	AAW71423	Aaw71423 Peptide h
32	47	34.3	21	5	ABG78944	Abg78944 Human bre
33	47	34.3	21	6	ABJ37799	Abj37799 Human tum
34	47	34.3	21	6	ABJ37767	Abj37767 Human tum
35	47	34.3	21	7	ADL93196	Adl93196 Human bre
36	47	34.3	21	7	ADL93249	Adl93249 Human bre
37	46	33.6	17	4	AAW74344	Aaw74344 Peptide C
38	45	32.8	28	4	AAU14021	Aau14021 Peptide s
39	45	32.8	28	6	ABO10240	Abol10240 Heptad re
40	45	32.8	28	7	ADB67071	Adb67071 Canonical
41	45	32.8	28	8	ADL99341	Adl99341 Nanostruc
42	44	32.1	24	3	AAV95963	Aay95963 Coiled-co
43	44	32.1	28	2	AAR31979	Aar31979 SSP4 poly
44	43	31.4	14	7	ADI39146	Adi39146 Amphipath
45	43	31.4	20	6	ABJ37808	Abj37808 Human tum

ALIGNMENTS

RESULT 1  
AAB74346  
ID AAB74346 standard; peptide; 28 AA.  
XX AC AAB74346;

DT 02-JUL-2001 (first entry)

DE Peptide SAF-p2B.

KW Atomic Force Microscopy; AFM.

OS Unidentified.

PN WO200121646-A1.

PD 29-MAR-2001.

PF 18-SEP-2000; 2000WO-GB003576.

PR 17-SEP-1999; 99GB-00022013.

PA (UYSU-) UNIV SUSSEX.

PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;

DR WPI; 2001-335468/35.

PT New protein structures with peptide monomer units, useful in Atomic Force Microscopy, purifying biological fluids, promoting tissue repair and tissue engineering, or constructing nanoscale molecular sieves.

PS Disclosure; Page 9; 45pp; English.

XX The present invention relates to a protein structure with several first peptide monomer units arranged in a first strand and several second peptide monomer units arranged in a second strand. The structure is useful in Atomic Force Microscopy and a number of other applications. The present sequence is the peptide monomer unit SAF-p2B

SQ Sequence 28 AA;

Query Match 100.0%; Score 137; DB 4; Length 28;

Best Local Similarity 100.0%; Pred. No. 3.2e-11;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIRALKWNAHLKQEIQAIEQIALEQ 28

|||||

Db 1 KIRALKKNAHLKOEIAALEQIEAALAEQ 28

RESULT 2

AAB74347

ID AAB74347 standard; peptide; 28 AA.

AC AAB74347;

XX 02-JUL-2001 (first entry)

XX Peptide SAF-p2C.

XX Atomic Force Microscopy; AFM.

XX Unidentified.

XX WO200121646-A1.

XX 29-MAR-2001.

XX 18-SEP-2000; 2000WO-GB003576.

XX 17-SEP-1999; 99GB-00022013.

XX (UYSU-) UNIV SUSSEX.

XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;

XX WPI; 2001-335468/35.

XX New protein structures with peptide monomer units, useful in Atomic Force

XX Microscopy, purifying biological fluids, promoting tissue repair and

XX tissue engineering, or constructing nanoscale molecular sieves.

XX Claim 16; Page 26; 45pp; English.

XX The present invention relates to a protein structure with several first

XX peptide monomer units arranged in a first strand and several second

XX peptide monomer units arranged in a second strand. The structure is

XX useful in Atomic Force Microscopy and a number of other applications. The

XX present sequence is the peptide monomer unit SAF-p2C

XX Sequence 28 AA;

Query Match 100.0%; Score 137; DB 4; Length 28;

Best Local Similarity 100.0%; Pred. No. 3.2e-11;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIRALKKNAHLKOEIAALEQIEAALAEQ 28

DB 1 KIRALKKNAHLKOEIAALEQIEAALAEQ 28

AC AAB74349;

XX 02-JUL-2001 (first entry)

XX Peptide SAF-p2E.

XX Atomic Force Microscopy; AFM.

XX Unidentified.

XX WO200121646-A1.

XX 29-MAR-2001.

XX 18-SEP-2000; 2000WO-GB003576.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX 17-SEP-1999; 99GB-00022013.

XX (UYSU-) UNIV SUSSEX.

XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;

XX WPI; 2001-335468/35.

XX New protein structures with peptide monomer units, useful in Atomic Force

XX Microscopy, purifying biological fluids, promoting tissue repair and

XX tissue engineering, or constructing nanoscale molecular sieves.

XX Disclosure; Page 9; 45pp; English.

XX The present invention relates to a protein structure with several first

XX peptide monomer units arranged in a first strand and several second

XX peptide monomer units arranged in a second strand. The structure is

XX useful in Atomic Force Microscopy and a number of other applications. The

XX present sequence is the peptide monomer unit SAF-p2E

XX Sequence 28 AA;

Query Match 90.5%; Score 124; DB 4; Length 28;

Best Local Similarity 96.4%; Pred. No. 1.7e-09;

Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIRALKKNAHLKOEIAALEQIEAALAEQ 28

DB 1 KIRALKKNAHLKOEIAALEQIEAALAEQ 28

RESULT 4

AAB74348

ID AAB74348 standard; peptide; 28 AA.

AC AAB74348;

XX 02-JUL-2001 (first entry)

XX Peptide SAF-p2D.

XX Atomic Force Microscopy; AFM.

XX Unidentified.

XX WO200121646-A1.

XX 29-MAR-2001.

XX 18-SEP-2000; 2000WO-GB003576.

XX 17-SEP-1999; 99GB-00022013.

XX (UYSU-) UNIV SUSSEX.

XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;

XX WPI; 2001-335468/35.

XX New protein structures with peptide monomer units, useful in Atomic Force

XX Microscopy, purifying biological fluids, promoting tissue repair and

XX tissue engineering, or constructing nanoscale molecular sieves.

XX Disclosure; Page 9; 45pp; English.

XX The present invention relates to a protein structure with several first

XX peptide monomer units arranged in a first strand and several second

XX peptide monomer units arranged in a second strand. The structure is

XX useful in Atomic Force Microscopy and a number of other applications. The

XX present sequence is the peptide monomer unit SAF-p2D

XX Sequence 28 AA;

Query Match 100.0%; Score 137; DB 4; Length 28;

Best Local Similarity 100.0%; Pred. No. 3.2e-11;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIRALKKNAHLKOEIAALEQIEAALAEQ 28

DB 1 KIRALKKNAHLKOEIAALEQIEAALAEQ 28

AC AAB74349;

XX 02-JUL-2001 (first entry)

XX Peptide SAF-p2E.

XX Atomic Force Microscopy; AFM.

XX Unidentified.

XX WO200121646-A1.

XX 29-MAR-2001.

XX 18-SEP-2000; 2000WO-GB003576.

XX 17-SEP-1999; 99GB-00022013.

XX (UYSU-) UNIV SUSSEX.

XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;

XX WPI; 2001-335468/35.

XX New protein structures with peptide monomer units, useful in Atomic Force

XX Microscopy, purifying biological fluids, promoting tissue repair and

XX tissue engineering, or constructing nanoscale molecular sieves.

XX Disclosure; Page 9; 45pp; English.

XX The present invention relates to a protein structure with several first

XX peptide monomer units arranged in a first strand and several second

XX peptide monomer units arranged in a second strand. The structure is

XX useful in Atomic Force Microscopy and a number of other applications. The

XX present sequence is the peptide monomer unit SAF-p2D

XX Sequence 28 AA;

```
Query Match      89.8%; Score 123; DB 4; Length 28;
Best Local Similarity 96.4%; Pred. No. 2.4e-09;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIRALKKQNAHLKQEIATALEQSIAALEQ 28
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1 KIRALKKQNAHLKQEIATALEQSIAALEQ 28

RESULT 5
AAB74356
ID AAB74356 standard; peptide; 28 AA.
XX AC AAB74356;
XX 02-JUL-2001 (first entry)
XX DE Peptide used to form blunt-ended heterodimers.
XX KW Atomic Force Microscopy; AFM.
XX OS Unidentified.
XX PN WO200121646-A1.
XX PD 29-MAR-2001.
XX PF 18-SEP-2000; 2000WO-GB003576.
XX PR 17-SEP-1999; 99GB-00022013.
XX PA (UYSU-) UNIV SUSSEX.
PI Woolfson DN, Walshaw J, Pandya MJ, Colyer J;
DR WPI; 2001-335468/35.
XX New protein structures with peptide monomer units, useful in Atomic Force
PT Microscopy, purifying biological fluids, promoting tissue repair and
PT tissue engineering, or constructing nanoscale molecular sieves.
XX Claim 16; Page 8; 45pp; English.
XX CC The present invention relates to a protein structure with several first
CC peptide monomer units arranged in a first strand and several second
CC peptide monomer units arranged in a second strand. The structure is
CC useful in Atomic Force Microscopy and a number of other applications. The
CC present sequence is a peptide used to form blunt-ended heterodimers
XX SQ Sequence 28 AA;

Query Match      89.8%; Score 123; DB 4; Length 28;
Best Local Similarity 96.4%; Pred. No. 2.4e-09;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIRALKKQNAHLKQEIATALEQSIAALEQ 28
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1 KIRALKKQNAHLKQEIATALEQSIAALEQ 28

RESULT 6
AAB74352
ID AAB74352 standard; peptide; 28 AA.
XX AC AAB74352;
XX 02-JUL-2001 (first entry)
XX DE Peptide SAF-p2.
XX KW Atomic Force Microscopy; AFM.
XX
```

```
OS Unidentified.
XX WO200121646-A1.
XX 29-MAR-2001.
XX 18-SEP-2000; 2000WO-GB003576.
XX 17-SEP-1999; 99GB-00022013.
XX (UYSU-) UNIV SUSSEX.
XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;
XX WPI; 2001-335468/35.
XX New protein structures with peptide monomer units, useful in Atomic Force
PT Microscopy, purifying biological fluids, promoting tissue repair and
PT tissue engineering, or constructing nanoscale molecular sieves.
XX Claim 16; Page 26; 45pp; English.
XX CC The present invention relates to a protein structure with several first
CC peptide monomer units arranged in a first strand and several second
CC peptide monomer units arranged in a second strand. The structure is
CC useful in Atomic Force Microscopy and a number of other applications. The
CC present sequence is the peptide monomer unit SAF-p2
XX SQ Sequence 28 AA;

Query Match      89.8%; Score 123; DB 4; Length 28;
Best Local Similarity 96.4%; Pred. No. 2.4e-09;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIRALKKQNAHLKQEIATALEQSIAALEQ 28
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1 KIRALKKQNAHLKQEIATALEQSIAALEQ 28

RESULT 7
AAB74345
ID AAB74345 standard; peptide; 28 AA.
XX AC AAB74345;
XX 02-JUL-2001 (first entry)
XX DE Peptide SAF-p2A.
XX KW Atomic Force Microscopy; AFM.
XX OS Unidentified.
XX PN WO200121646-A1.
XX PD 29-MAR-2001.
XX PF 18-SEP-2000; 2000WO-GB003576.
XX PR 17-SEP-1999; 99GB-00022013.
XX PA (UYSU-) UNIV SUSSEX.
XX Woolfson DN, Walshaw J, Pandya MJ, Colyer J;
XX WPI; 2001-335468/35.
XX New protein structures with peptide monomer units, useful in Atomic Force
PT Microscopy, purifying biological fluids, promoting tissue repair and
PT tissue engineering, or constructing nanoscale molecular sieves.
XX Disclosure; Page 9; 45pp; English.
XX
```





New protein structures with peptide monomer units, useful in Atomic Force Microscopy, purifying biological fluids, promoting tissue repair and

PT tissue engineering, or constructing nanoscale molecular sieves.

XX

PS

XX

XX

CC

CC

CC

CC

CC

XX

SQ

Disclosure; Page 9; 45pp; English.

The present invention relates to a protein structure with several first peptide monomer units arranged in a first strand and several second peptide monomer units arranged in a second strand. The structure is useful in Atomic Force Microscopy and a number of other applications. The present sequence is the peptide monomer unit SAF-plB

Sequence 28 AA;

Query Match

Best Local Similarity 54.0%; Score 74; DB 4; Length 28;

Matches 20; Conservative 71.4%; Pred. No. 0.0082;

Mismatches 8; Indels 0; Gaps 0;

OY 1 KIRALKKNAHLKQETIALEQ 28

DB 1 KIRALKKNAHLKQETIALEQ 28

Search completed: April 27, 2005, 15:02:16

Job time : 123 secs

**BEST AVAILABLE COPY**

**THIS PAGE LEFT BLANK**



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	36	26.3	25	2	S39391	calpain II heavy c	
2	32	23.4	23	2	S23518	stromelysin (EC 3.	
3	31	22.6	16	2	E37290	homeotic protein G	
4	31	22.6	21	2	I65270	collagen alpha 1(I	
5	30	21.9	22	2	B35372	unidentified low M	
6	29	21.2	24	2	A33262	heparin-binding gr	
7	29	21.2	24	2	I38253	T-cell acute lymph	
8	28.5	20.8	20	2	A53592	H+-exporting ATPas	
9	28	20.4	19	2	S04955	hemocyanin chain I	
10	28	20.4	24	2	C47689	flagellar core pro	
11	28	20.4	27	2	G61002	transformation-sen	
12	27	19.7	21	2	A59325	probable bacteriop	
13	26.5	19.3	27	2	PQ0844	DNA-binding protei	
14	26	19.0	15	2	PA0036	glycine cleavage s	
15	26	19.0	19	2	FN0467	nitrogenase (EC 1.	
16	26	19.0	20	2	B46236	transferrin prote	
17	26	19.0	25	2	B47689	flagellar core pro	
18	25	18.2	10	2	A43405	6-phosphofructo-2-	
19	25	18.2	20	2	S35970	ribosomal protein	
20	25	18.2	21	2	S35978	ribosomal protein	
21	25	18.2	23	2	S35975	ribosomal protein	
22	25	18.2	23	2	S35976	ribosomal protein	
23	25	18.2	24	2	J46513	troponin I - rabbi	
24	25	18.2	25	1	IC4278	ribosomal protein	
25	25	18.2	25	2	J01617	ribosomal protein	
26	25	18.2	25	2	JC4685	ribosomal protein	
27	25	18.2	25	2	S38425	ribosomal protein	
28	25	18.2	25	2	T49214	ribosomal protein	
29	25	18.2	25	2	T06233	ribosomal protein	

```

RESULT 3
E37290
homeotic protein Gsh-5 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jan-1992 #sequence_revision 17-Jan-1992 #text_change 16-Aug-2004
C:Accession: E37290; E38809
R:Singh, G.; Kaur, S.; Stock, J.L.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Potter
Proc. Natl. Acad. Sci. U.S.A. 88, 10706-10710, 1991
A:Title: Identification of 10 murine homeobox genes.
A:Reference number: A37290; MUID:92073356; PMID:1683707
A:Accession: E37290
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-16 <SIN>
C:Superfamily: homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:1-14/Domain: homeobox homology (fragment) <HOX>

Query Match 22.6%; Score 31; DB 2; Length 16;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RALKWKNA 10
DB 8 RRMKWRNS 15

RESULT 4
I65270
collagen alpha 1(I) chain - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C:Accession: I65270
R:Genovese, C.; Rowe, D.; Kream, B.
Biochemistry 23, 6210-6216, 1984
A:Title: Construction of DNA sequences complementary to rat alpha-1 and alpha-2 collagen
A:Reference number: I52392; MUID:85122694; PMID:6395893
A:Accession: I65270
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-21 <RES>
A:Cross-references: UNIPROT:Q63076; GB:M12200; NID:G203191; PID:9203195
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 22.6%; Score 31; DB 2; Length 21;
Best Local Similarity 50.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 11 HLKQEIATAEQE 22
DB 8 HCKNSIAYLDEE 19

RESULT 5
B35372
identified low M-r protein - Yersinia pseudotuberculosis (fragment)
C:Species: Yersinia pseudotuberculosis
C:Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 09-Jul-2004
C:Accession: B35372
R:Han, O.; Miller, V.P.; Liu, H.
J. Biol. Chem. 265, 8033-8041, 1990
A:Title: Mechanistic studies of the biosynthesis of 3,6-dideoxyhexoses in Yersinia pseud
H: dichlorophenolindolphenol oxidoreductase activity.
A:Reference number: A35372; MUID:90243678; PMID:2159466
A:Accession: B35372
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-22 <HAN>
A:Cross-references: UNIPROT:Q7MON2

Query Match 21.9%; Score 30; DB 2; Length 22;
Best Local Similarity 70.0%; Pred. No. 8.7e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 17 AALEQRIAL 26
DB 6 AALVKELAL 15

RESULT 6
A33262
heparin-binding growth factor A light chain - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 09-Jul-2004
C:Accession: A33262
R:Zarnegar, R.; Muga, S.; Enghild, J.; Michalopoulos, G.
Biochem. Biophys. Res. Commun. 163, 1370-1376, 1989
A:Title: NH-2-terminal amino acid sequence of rabbit hepatopoietin A, a heparin-binding
A:Reference number: A33262; MUID:8932048; PMID:2528955
A:Accession: A33262
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <ZAR>
A:Cross-references: UNIPROT:P13571
C:Superfamily: hepatocyte growth factor/macrophage stimulating protein 1; kringle homolog
C:Keywords: growth factor; heparin binding; kringle

Query Match 21.2%; Score 29; DB 2; Length 24;
Best Local Similarity 44.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 ALKWKNAHL 12
DB 16 SLKYNKXHI 24

RESULT 7
I38253
T-cell acute lymphocytic leukemia 1 protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999
C:Accession: I38253
R:Chen, Q.; Yang, C.Y.; Tsan, J.T.; Xia, Y.; Ragab, A.H.; Peiper, S.C.; Carroll, A.; Bae
J. Exp. Med. 172, 1403-1408, 1990
A:Title: Coding sequences of the tal-1 gene are disrupted by chromosome translocation in
A:Reference number: I38253; MUID:91037802; PMID:2230650
A:Accession: I38253
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-24 <RES>
A:Cross-references: EMBL:X58621; NID:G36683; PIDN:CAA41476.1; PID:G36684
C:Genetics:
A:Gene: GDB:TAL1; SCL; TCL5
A:Cross-references: GDB:I20759; OMIM:187040
A:Map position: lp32-lp32

Query Match 21.2%; Score 29; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 RALKWKNAH 11
DB 10 RASKWPHAY 18

RESULT 8
A53592
Hx-exporting ATPase (EC 3.6.3.6) 24k chain - Thermus aquaticus (fragment)
C:Species: Thermus aquaticus
C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C:Accession: A53592
R:Yokoyama, K.; Akabane, Y.; Ishii, N.; Yoshida, M.
J. Biol. Chem. 269, 12248-12253, 1994
A:Title: Isolation of prokaryotic V-0V-1-ATPase from a thermophilic eubacterium Thermus
A:Reference number: A53592; MUID:94216345; PMID:8163530
A:Accession: A53592

```

A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <YOK>  
A:Cross-references: UNIPROT:Q7M196  
C:Keywords: hydrolase

Query Match 20.8%; Score 28.5; DB 2; Length 20;  
Best Local Similarity 52.6%; Pred. No. 1.2e+03;  
Matches 10; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

QY 10 AHLKQEIQAALAEQ 28  
| | | | | | | | | | | | | | | | | | | | | |  
Db 5 AILSQEV---EAEIQALLQ 20

## RESULT 9

S00495  
hemocyanin chain III - Japanese spiny lobster (fragment)

C:Species: Panulirus japonicus (Japanese spiny lobster)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
A:Accession: S00495

R:Makino, N.; Kimura, S.

A:Title: Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.

A:Reference number: S00492; MUID:88196131; PMID:3360019

A:Accession: S00495

A:Molecule type: protein

A:Residues: 1-19 <NAK>

A:Cross-references: UNIPROT:P82313

C:Superfamily: hemocyanin

C:Keywords: copper; hemolymph; hexamer; oxygen carrier

Query Match 20.4%; Score 28; DB 2; Length 19;  
Best Local Similarity 62.5%; Pred. No. 1.4e+03;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 NAHLKQEI 16

||| :| :|  
Db 7 NAHKQDI 14

## RESULT 10

C47689

flagellar core protein, 34K - Treponema hyodysenteriae (fragment)

C:Species: Treponema hyodysenteriae

C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

A:Accession: C47689

R:Koopman, M.B.; Baate, E.; van Vorstenbosch, C.J.; van der Zeijst, B.A.; Kusters, J.G.

J. Gen. Microbiol. 138, 2697-2706, 1992

A:Title: The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are composed of

A:Reference number: A47689; MUID:93139764; PMID:1487733

A:Contents: C5: Treponema

A:Accession: C47689

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-24 <KOO>

A:Cross-references: UNIPROT:Q7M132

A>Note: sequence extracted from NCBI backbone (NCBIP:123402)

Query Match 20.4%; Score 28; DB 2; Length 24;  
Best Local Similarity 45.5%; Pred. No. 1.7e+03;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 RALKWNAHLK 13

| | | | | | | | | | | | | | | | | | | | | |  
Db 14 RTLKFRNVDLR 24

## RESULT 11

G61002

transformation-sensitive protein IEF-3613 - human (fragments)

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999

C:Accession: G61002

R:Bauw, G.; Raemussen, H.H.; Van Den Bulcke, M.; Van Damme, J.; Puype, M.; Gesser, B.;

Electrophoresis 11, 528-536, 1990

A:Title: Two-dimensional gel electrophoresis, protein electrophoresis and microsequencing

A:Reference number: A61002; MUID:91031404; PMID:1699755

A:Accession: G61002

A:Molecule type: protein

A:Residues: 1-27 <BAU>

A>Note: this material corresponds to transformed epithelial amnion cell (AMA) database

Query Match 20.4%; Score 28; DB 2; Length 27;  
Best Local Similarity 31.2%; Pred. No. 2e+03;  
Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 8 KNAHLKQEIQAALAEQEI 23

| | | | | | | | | | | | | | | | | | | | | |  
Db 6 KLAYINPDLLALEETL 21

## RESULT 12

A59325

probable bacteriophage receptor BactA [imported] - Bacteroides fragilis (fragment)

C:Species: Bacteroides fragilis

C:Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000 #text\_change 09-Jul-2004

A:Accession: A59325

R:Frias-Lopez, J.

submitted to the Protein Sequence Database, July 2000

A:Description: Identification of cell wall proteins of Bacteroides fragilis to which bac

A:Reference number: A59325

A:Accession: A59325

A:Molecule type: protein

A:Residues: 1-21 <FRI>

A:Cross-references: UNIPROT:Q7M123

A:Experimental source: strain ATCC 51477

A>Note: putative receptor for bacteriophage B40-8

Query Match 19.7%; Score 27; DB 2; Length 21;  
Best Local Similarity 58.3%; Pred. No. 2.1e+03;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 EIALAEQEIQAAL 26

| | | | | | | | | | | | | | | | | | | | | |  
Db 8 EPAALHDEMLAL 19

## RESULT 13

PQ0844

DNA-binding protein AcBBP1 - Azorhizobium caulinodans (fragment)

C:Species: Azorhizobium caulinodans

C:Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 09-Jul-2004

A:Accession: PQ0844

R:Welters, P.; Metz, B.; Felix, G.; Palme, K.; Szczyglowski, K.; de Bruijn, P.J.

Plant Physiol. 102, 1095-1107, 1993

A:Title: Interaction of a rhizobial DNA-binding protein with the promoter region of a pl

A:Reference number: PQ0844; MUID:94105338; PMID:8278541

A:Accession: PQ0844

A:Molecule type: protein

A:Residues: 1-27 <WEL>

A:Cross-references: UNIPROT:Q7M1A8

A:Experimental source: strain ORS571

Query Match 19.3%; Score 26.5; DB 2; Length 27;  
Best Local Similarity 30.4%; Pred. No. 3.1e+03;  
Matches 7; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 2 IRALKWKN-AHLKQEIQAALAEQEI 23

| | | | | | | | | | | | | | | | | | | | | |  
Db 3 MKLIVGRNFARLQEKGLTQSDV 25

## RESULT 14

PA0036

Thu Apr 28 07:26:23 2005

glycine cleavage system protein H1 and H2 - Arabidopsis thaliana (fragment)  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C:Accession: PA0036  
 R:Kamo, M.; Kawakami, T.; Miyatake, N.; Teugita, A.  
 submitted to JPIID, July 1994  
 A:Description: Separation and characterization of Arabidopsis proteins by two-dimensional gel electrophoresis  
 A:Reference number: PA0001  
 A:Accession: PA0036  
 A:Molecule type: protein  
 A:Residues: 1-15 <KAM>  
 A:Cross-references: UNIPROT:Q7MLX0  
 A:Experimental source: stem

Query Match 19.0%; Score 26; DB 2; Length 15;  
 Best Local Similarity 40.0%; Pred. No. 2e+03;  
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 IRALKKNAH 11  
 : ||| :  
 DB 4 LEGKYANCH 13

## RESULT 15

PN0467  
 nitrogenase (EC 1.18.6.1) molybdenum-iron protein alpha chain - Azotobacter chroococcum  
 C:Species: Azotobacter chroococcum  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: PN0467  
 R:Jones, R.; Woodley, P.; Birkmann-Zinoni, A.; Robson, R.L.  
 Gene 123, 145-146, 1993  
 A:Title: The nifH gene encoding the Fe protein Component of the molybdenum nitrogenase  
 A:Reference number: JN0516; MUID:93138425; PMID:8423000  
 A:Accession: PN0467  
 A:Molecule type: DNA  
 A:Residues: 1-19 <JON>  
 A:Cross-references: UNIPROT:Q43964; GB:M73020; NID:gi42326; PIDN:AAA22141.1; PID:G289238  
 C:Genetics:  
 A:Gene: nifD  
 C:Superfamily: dinitrogenase alpha chain; nitrogenase vanadium-iron protein alpha chain  
 C:Keywords: ATP; nitrogen fixation; oxidoreductase

Query Match 19.0%; Score 26; DB 2; Length 19;  
 Best Local Similarity 36.4%; Pred. No. 2.5e+03;  
 Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 13 KQEIAPAEQEI 23  
 :||: |||:  
 DB 6 REEVESLIQEV 16

Search completed: April 27, 2005, 15:07:13  
 Job time : 29.5 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	36	26.3	25	2	Q7M386	Q7m386 sus scrofa
2	34	24.8	27	2	Q661K4	Q661k4 borrelia gar
3	33	24.1	24	2	Q9RP49	Q9rp49 mycobacteri
4	32	23.4	25	2	Q7QXH6	Q7qxh6 giardia lam
5	32	23.4	26	2	Q9TJ2S3	Q9tj2s3 solanum tub
6	31	22.6	16	2	Q9QW74	Q9qw74 mus sp. hom
7	31	22.6	21	2	Q63076	Q63076 rattus norv
8	31	22.6	22	2	Q9H002	Q9h002 homo sapien
9	30	21.9	22	2	Q7M0N2	Q7m0n2 yersinia pes
10	30	21.9	28	2	Q9IN95	Q9in95 human immun
11	29.5	21.5	28	2	Q6V4K8	Q6v4k8 streptococc
12	29	21.2	13	2	Q50038	Q50038 mycobacteri
13	29	21.2	23	2	Q6V106	Q6v106 mus spretus
14	29	21.2	24	1	HPTA.RABIT	P13571 oryctolagus
15	29	21.2	26	2	Q6BDL1	Q6bdl1 tricholoma
16	28.5	20.8	20	2	Q7M196	Q7m196 thermus aqu
17	28	20.4	19	1	HCY3.PANJA	P82313 panulirus j
18	28	20.4	21	2	Q9UGJ6	Q9ugj6 homo sapien
19	28	20.4	23	1	NIFD.ANANSL	P33177 anobaena sp
20	28	20.4	24	2	Q7M132	Q7m132 treponema h
21	28	20.4	25	1	LYTI.HOGCA	P61507 hogma carol
22	27.5	20.1	27	2	Q8T3I7	Q8t3i7 drosophila
23	27	19.7	19	2	Q9BPDG4	Q9bdg4 bos taurus
24	27	19.7	21	2	Q7M123	Q7m123 bacteroides
25	27	19.7	22	2	O44710	O44710 drosophila
26	27	19.7	23	2	Q94XD3	Q94xd3 pteroxanium
27	27	19.7	23	2	Q9AJQ7	Q9ajq7 rattus norv
28	27	19.7	24	2	Q9R1M2	Q9rim2 rattus norv
29	27	19.7	26	2	Q9NJR1	Q9njr1 euprymna ec
30	26.5	19.3	27	2	Q7M1A8	Q7mia8 azorhizobiu
31	26	19.0	13	2	Q9QW45	Q9qw45 rattus sp.

Best Local Similarity 52.9%; Pred. No. 2.2e+03; Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 KIRALKKWNHLKQIEA 17  
Db 4 KILALKKKNLNYKINIS 20

RESULT 3  
Q9RP49 Q9RP49 PRELIMINARY; PRT; 24 AA.  
AC Q9RP49; 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE Putative glycogen phosphorylase Glgp (Fragment).  
GN Name=glgp;  
OS Mycobacterium smegmatis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1772;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=mc2155;  
RX MEDLINE=20011333; PubMed=10542168;  
RA Belanger A.E., Hatfull G.F.;  
RT "Exponential-phase glycogen recycling is essential for growth of  
RT Mycobacterium smegmatis."  
RL J. Bacteriol. 181:6670-6678 (1999).  
DR EMBL; AF172946; AAF07897.1; --  
FT NON\_TER 24 24  
SQ SEQUENCE 24 AA; 2865 MW; 0DBICFD072F8A36B CRC64;

Query Match 24.1%; Score 33; DB 2; Length 24;  
Best Local Similarity 54.5%; Pred. No. 2.6e+03; Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 HLKQIEIALEQ 21  
Db 12 HLPERLAALER 22

RESULT 4  
Q7QXH6 Q7QXH6 PRELIMINARY; PRT; 25 AA.  
AC Q7QXH6; 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE GLP 14 5818 5741.  
OS Giardia lamblia ATCC 50803.  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.  
OX NCBI\_TaxID=184922;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WB CS;  
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,  
RA Olsen G.J., Sogin M.L.;  
RT "Draft sequence of the Giardia lamblia genome."  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AACB0100060; EAA39724.1; --  
DR GO; GO:0005840; C:ribosome; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
DR InterPro; IPR007836; Ribosomal L41.  
DR Pfam; PF05162; Ribosomal L41; 1  
SQ SEQUENCE 25 AA; 3267 MW; 9A0629C253733C4E CRC64;

Query Match 23.4%; Score 32; DB 2; Length 25;  
Best Local Similarity 46.2%; Pred. No. 3.7e+03;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KIRALKKWNHLK 13  
Db 9 RMRKLRKRHRMK 21

RESULT 5  
Q9T2S3 Q9T2S3 PRELIMINARY; PRT; 26 AA.  
AC Q9T2S3; 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Cytochrome-C reductase 55 kDa subunit (EC 1.10.2.2) (Fragment).  
OS Solanum tuberosum (potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94198758; PubMed=7764624;  
RA Braun H.P., Kruft V., Schmitz U.K.;  
RL Planta 193:99-106 (1994).  
DR MEROPS; M16.003; --  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0008121; F:ubiquinol-cytochrome-c reductase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001431; Insulinase like.  
DR Pfam; PF00675; Peptidase M16; 1.  
SQ SEQUENCE 26 AA; 3042 MW; C92C5787C8F12AC8 CRC64;

Query Match 23.4%; Score 32; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 3.9e+03; Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 13 KOEIAALEOEI 23  
Db 1 KRPIRALEEEI 11

RESULT 6  
Q9QW74 Q9QW74 PRELIMINARY; PRT; 16 AA.  
AC Q9QW74; 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Homeobox GSH-5 (Fragment).  
OS Mus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10095;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92073356; PubMed=1683707;  
RA Singh G., Kaur S., Stock S.J., Jenkins N.A., Gilbert D.J.,  
RA Copeland N.G., Potter S.S.;  
RT "Identification of 10 murine homeobox genes."  
RL Proc. Natl. Acad. Sci. U.S.A. 88:10706-10710 (1991).  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeodomain-like.  
DR Prodom; PD000010; Homeobox; 1.  
DR PROSITE; PS00071; Homeobox\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1 16  
FT NON\_TER 16 16

SQ SEQUENCE 16 AA; 2178 MW; AEC306578EE24361 CRC64;  
Query Match 22.6%; Score 31; DB 2; Length 16;  
Best Local Similarity 50.0%; Pred. No. 3.1e+03;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 3 RALKWKNA 10  
| :||:|  
Db 8 REMKWRNS 15  
| :||:|  
RESULT 7  
Q63076 PRELIMINARY; PRT; 21 AA.  
AC Q63076;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Alpha-2 type I collagen (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal calvaria;  
RX MEDLINE=85122694; PubMed=6395893;  
RA Genovese C., Rowe D., Kream B.;  
RT "Construction of DNA sequences complementary to rat alpha 1 and alpha  
RT 2 collagen mRNA and their use in studying the regulation of type I  
RT collagen synthesis by 1,25-dihydroxyvitamin D."  
RL Biochemistry 23:6210-6216(1984).  
DR EMBL; M12200; AAA40835.1; -.  
DR PIR; I65270; I65270  
DR GO; GO:0005581; C:collagen; IEA.  
DR GO; GO:0005201; F:extracellular matrix structural constituent; IEA.  
DR InterPro; IPR000885; Fib\_collagen\_C.  
DR Pfam; PF01410; COLFI; 1.  
DR ProDom; PD002078; Fib\_collagen\_C; 1.  
KW Collagen.  
FT NON\_TER 1 1  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2358 MW; B86AGECCCEE46D60 CRC64;  
Query Match 22.6%; Score 31; DB 2; Length 21;  
Best Local Similarity 50.0%; Pred. No. 4.2e+03;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 11 HLKQEIAALEQE 22  
| :||:|  
Db 8 HCKNSIAYLDEE 19  
| :||:|  
RESULT 8  
Q9H002 PRELIMINARY; PRT; 22 AA.  
AC Q9H002;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ATP-binding cassette transporter-1 (Fragment).  
GN Name=ABC-1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Porach-Oezcuernuez M., Langmann T., Schmitz G.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Porach-Oezcuernuez M.K.;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ252277; CAC21428.1; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
KW ATP-binding.  
FT NON\_TER 22 22  
SQ SEQUENCE 22 AA; 2831 MW; 6183559EA0C22B53 CRC64;  
Query Match 22.6%; Score 31; DB 2; Length 22;  
Best Local Similarity 55.6%; Pred. No. 4.4e+03;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KIRALKWQN 9  
| :||:|  
Db 6 QLRLLLWKN 14  
| :||:|  
RESULT 9  
Q7MON2 PRELIMINARY; PRT; 22 AA.  
AC Q7MON2;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Unidentified low M-r protein (Fragment).  
OS Versinia pseudotuberculosis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Versinia.  
OX NCBI\_TaxID=633;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=90243678; PubMed=2159466;  
RA Han O., Miller V.P., Liu H.;  
RT "Mechanistic studies of the biosynthesis of 3,6-dideoxyhexoses in  
RT Versinia pseudotuberculosis. Purification and characterization of CDP-  
RT 6-deoxy-delta(3,4)-glucose reductase based on its  
RT NADH:chlorophenolindolphenol oxidoreductase activity."  
RL J. Biol. Chem. 265:8033-8041(1990).  
DR PIR; B35372; B35372.  
DR NON\_TER 22 22  
SQ SEQUENCE 22 AA; 2192 MW; 3E88CD8AAC8680BB CRC64;  
Query Match 21.9%; Score 30; DB 2; Length 22;  
Best Local Similarity 70.0%; Pred. No. 5.9e+03;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 17 AALSQEIAAL 26  
| :||:|  
Db 6 AALVKELAAL 15  
| :||:|  
RESULT 10  
Q9IN95 PRELIMINARY; PRT; 28 AA.  
AC Q9IN95;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Gag protein (Fragment).  
GN Name=gag;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99260287; PubMed=10331444; DOI=10.1089/089922299310953;  
RA Mokili J.L., Wade C.M., Burns S.M., Cutting W.A., Bopopi J.M.,  
RA Green S.D., Featherer J.F., Simmonds P.;  
RT "Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural  
RT Democratic Republic of Congo".  
RL AIDS Res. Hum. Retroviruses 15:655-664(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Mokili J.L.K.;

RL Submitted (APR-1999) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AF144854; AAF69080.1; -  
FT NON\_TER 1  
SQ SEQUENCE 28 AA; 3322 MW; 76749DCB32968AA9 CRC64;

Query Match 21.9%; Score 30; DB 2; Length 28;  
Best Local Similarity 66.7%; Pred. No. 7.6e+03;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 WKNAHL 12  
|:|:|  
Db 2 WKDSHL 7

## RESULT 11

Q6V4K8 PRELIMINARY; PRT; 28 AA.  
AC Q6V4K8;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE M protein (Fragment).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]\_TaxID=1314;  
RP SEQUENCE FROM N.A.  
RC STRAIN=NS696;  
RX PubMed=14688117; DOI=10.1128/JAI.72.1.364-370.2004;  
RA McKay F.C., McArthur J.D., Sanderson-Smith M.L., Gardam S.,  
RA Currie B.J., Sriprakash K.S., Fagan P.K., Towers R.J., Batzloff M.R.,  
RA Chhatwal G.S., Ranson M., Walker M.J.;  
RT "Plasminogen binding by group A streptococcal isolates from a region  
of hypervirulence for streptococcal skin infection and a high  
incidence of invasive infection."  
RL Infect. Immun. 72:364-370(2004).  
DR EMBL; AY351858; AAQ64528.1; -  
FT NON\_TER 1  
FT NON\_TER 28  
SQ SEQUENCE 28 AA; 3463 MW; 013B1C94B01B96ED CRC64;

Query Match 21.5%; Score 29.5; DB 2; Length 28;  
Best Local Similarity 38.1%; Pred. No. 8.8e+03;  
Matches 8; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 7 WKNAHLKOEIAALEQIEALE 27  
|:|:|:|:|  
Db 1 WDRQRLEKE---LEEKKEALE 18

## RESULT 12

Q50038 PRELIMINARY; PRT; 13 AA.  
AC Q50038;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-NOV-1996 (TReMBLrel. 01, Last annotation update)  
DE U22661.  
OS Mycobacterium leprae.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]\_TaxID=1769;  
RP SEQUENCE FROM N.A.  
RA Smith D.R.;  
RL Submitted (APR-1995) to the EMBL/GenBank/DBSJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Robison K.;  
RL Submitted (SEP-1994) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; U15182; AAA62970.1; -  
SQ SEQUENCE 13 AA; 1589 MW; C7CCD7E47D025B06 CRC64;

Query Match 21.2%; Score 29; DB 2; Length 13;  
Best Local Similarity 45.5%; Pred. No. 4.6e+03;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 HLKQEIAALEQ 21  
|:|:|:|  
Db 3 HVQOEYLGVEQ 13

## RESULT 13

Q6V106 PRELIMINARY; PRT; 23 AA.  
AC Q6V106;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE Heme oxygenase 1 (Fragment).  
GN Name=Hol;  
OS Mus spretus (Western wild mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Cabrera-Lague J.M., Prieto-Alamo M.J., Pueyo C.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AY355145; AAQ55831.1; -  
FT NON\_TER 1  
FT NON\_TER 23  
SQ SEQUENCE 23 AA; 2852 MW; F8D0C3F6975C5900 CRC64;

Query Match 21.2%; Score 29; DB 2; Length 23;  
Best Local Similarity 75.0%; Pred. No. 8.3e+03;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 17 AALEQEIA 24  
|:|:|:|:|  
Db 4 AALEQDMA 11

## RESULT 14

HPTA\_RABIT STANDARD; PRT; 24 AA.  
AC P13571;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Heptapoeitin A light chain (HPTA) (Fragment).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=89392048; PubMed=2528955;  
RA Zarnegar R., Muga S., Enghild J., Michalopoulos G.;  
RT "NH2-terminal amino acid sequence of rabbit heptapoeitin A, a  
heparin-binding polypeptide growth factor for hepatocytes."  
RL Biochem. Biophys. Res. Commun. 163:1370-1376(1989).  
CC -!- FUNCTION: HPTA is an acidic heparin-binding growth factor for  
hepatocytes.  
CC -!- SUBUNIT: Heterodimer of an heavy and a light chain linked by  
disulfide bond(s).  
DR PIR; A33262; A33262.  
KW Direct protein sequencing; Growth factor; Heparin-binding.  
FT NON\_TER 24  
FT NON\_TER 24  
SQ SEQUENCE 24 AA; 2768 MW; E78BF9B807262FE CRC64;

Query Match 21.2%; Score 29; DB 1; Length 24;  
Best Local Similarity 44.4%; Pred. No. 8.7e+03;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;



QY 4 ALKWKNAHL 12  
:|||||:  
Db 16 SLKYRNKHI 24

RESULT 15  
Q6BDL1 PRELIMINARY; PRT; 26 AA.  
AC Q6BDL1;  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Hypothetical protein (Fragment).  
OS Tricholoma matsutake.  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
OC Agaricales; Tricholomataceae; Tricholoma.  
OX NCBI\_TaxID=40145;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Murata H.;  
RT "Characterization of the insertion sites of marY1, the gypsy-type  
retrotransposon from the ectomycorrhizal basidiomycete Tricholoma  
matsutake strain Y1, in the genome the fungus based on the inter-  
retrotransposon amplified polymorphism analysis.";  
RL Submitted (JAN-2004) to the EMBL/GenBank/DDJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; AB160891; BAD32667.1; -;  
DR GO; GO:0000785; C:chromatin; IEA.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003682; F:chromatin binding; IEA.  
DR GO; GO:0006333; P:chromatin assembly or disassembly; IEA.  
DR InterPro; IPR000953; Chromo.  
DR Pfam; PF00385; Chromo; 1.  
DR PROSITE; PS00598; CHROMO\_1; UNKNOWN\_1.  
DR PROSITE; PSS0013; CHROMO\_2; 1.  
KW Hypothetical protein; Nuclear protein.  
FT NON\_TER 1  
FT NON\_TER 26  
SQ SEQUENCE 26 AA; 3272 MW; 3E4B000853591617 CRC64;

Query Match 21.2%; Score 29; DB 2; Length 26;  
Best Local Similarity 80.0%; Pred. No. 9.4e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 LKWKV 9  
:|||||:  
Db 7 LKWKV 11

Search completed: April 27, 2005, 15:06:18  
Job time : 114 secs

**BEST AVAILABLE COPY**

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2005, 14:58:17 ; Search time 29.5 Seconds  
(without alignments)  
70.853 Million cell updates/sec

Title: US-10-088-417A-4  
Perfect score: 137  
Sequence: 1 KIRALKWNAHLKQEIATLAEQ 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 228236

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS-COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	34.3	21	4	US-09-834-759-534
2	45	32.8	26	4	US-09-403-861A-8
3	45	32.8	27	2	US-07-684-965-9
4	45	32.8	27	3	US-09-252-658-9
5	45	32.8	28	3	US-08-486-099-84
6	45	32.8	28	3	US-08-360-107A-94
7	45	32.8	28	3	US-08-484-223B-84
8	45	32.8	28	3	US-08-519-597-84
9	45	32.8	28	3	US-08-475-668A-84
10	45	32.8	28	3	US-08-485-551A-84
11	45	32.8	28	3	US-08-471-913A-84
12	45	32.8	28	3	US-08-485-264A-84
13	45	32.8	28	3	US-08-474-349A-84
14	45	32.8	28	4	US-08-255-208A-20
15	45	32.8	28	4	US-08-470-896-84
16	45	32.8	28	4	US-09-914-259-1
17	45	32.8	28	4	US-08-485-546A-84
18	45	32.8	28	4	US-09-350-841A-1549
19	45	32.8	28	4	US-09-350-841A-1566
20	45	32.8	28	4	US-09-350-841A-1573
21	45	32.8	28	4	US-08-487-266A-84
22	44	32.1	24	4	US-09-491-614B-29
23	44	32.1	28	1	PCT-US92-06412-1
24	44	32.1	28	5	US-08-182-175A-1
25	43	31.4	24	2	US-08-491-527A-12
26	42	30.7	24	2	US-08-491-527A-11
27	42	30.7	25	4	US-09-491-614B-25

28	42	30.7	26	2	US-08-690-011A-37	Sequence 37, Appl
29	42	30.7	26	3	US-09-299-495F-37	Sequence 37, Appl
30	40	29.2	14	2	US-08-491-527A-8	Sequence 8, Appl
31	39	28.5	22	3	US-09-248-588-98	Sequence 98, Appl
32	38	27.7	14	2	US-08-491-527A-9	Sequence 9, Appl
33	38	27.7	25	2	US-08-690-011A-2	Sequence 2, Appl
34	38	27.7	25	3	US-09-299-495F-2	Sequence 2, Appl
35	38	27.7	26	2	US-08-690-011A-3	Sequence 3, Appl
36	38	27.7	26	3	US-09-299-495F-3	Sequence 3, Appl
37	38	27.7	28	1	US-08-182-175A-3	Sequence 3, Appl
38	38	27.7	28	1	US-08-474-633A-70	Sequence 70, Appl
39	38	27.7	28	4	US-08-823-771-70	Sequence 70, Appl
40	38	27.7	28	5	PCT-US92-06412-3	Sequence 3, Appl
41	37	27.0	27	2	US-07-684-965-8	Sequence 8, Appl
42	37	27.0	27	3	US-09-252-658-8	Sequence 8, Appl
43	37	27.0	28	1	US-07-977-630-53	Sequence 53, Appl
44	36	26.3	14	2	US-08-491-527A-7	Sequence 7, Appl
45	36	26.3	14	2	US-08-491-527A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-834-759-534  
; Sequence 534, Application US/09834759  
; Patent No. 6680197  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.470C9  
; CURRENT APPLICATION NUMBER: US/09/834,759  
; CURRENT FILING DATE: 2001-04-13  
; NUMBER OF SEQ ID NOS: 547  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 534  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-759-534

Query Match 34.3%; Score 47; DB 4; Length 21;  
Best Local Similarity 57.9%; Pred. NO. 5.8;  
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 NAHLKQEIATLAEQIAALE 27  
DB 1 NCMLKKEIATMLKLEIATLK 19

RESULT 2  
US-09-403-861A-8  
; Sequence 8, Application US/09403861A  
; Patent No. 6833348  
; GENERAL INFORMATION:  
; APPLICANT: RICCARDI, Carlo  
; TITLE OF INVENTION: INTRACELLULAR MODULATORS OF APOPTIC CELL  
; DEATH PATHWAYS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001

Qy  
1 KIRALKWKNHLKQEIAALEQ 21

```

Db      4 KVEELSKNYHLENEVARLKK 24

RESULT 5
US-08-486-099-84
; Sequence 84, Application US/08486099
; Patent No. 6013263
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; TITLE OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,099
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-486-099-84

Query Match      32.8%; Score 45; DB 3; Length 28;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy      1 KIRALKWQVHLKQETAALEQ 21
; : | | | | | | | | | | | | | | | |
Db      7 KVEELSKNYHLENEVARLKK 27

RESULT 6
US-08-360-107A-94
; Sequence 94, Application US/08360107A
; Patent No. 6017536
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,107A
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-360-107A-94

Query Match      32.8%; Score 45; DB 3; Length 28;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy      1 KIRALKWQVHLKQETAALEQ 21
; : | | | | | | | | | | | | | | | |
Db      7 KVEELSKNYHLENEVARLKK 27

RESULT 7
US-08-484-223B-84
; Sequence 84, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/484,223B  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7872-029  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEFAX: (212) 869-9741/8864  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 84:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 28 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: peptide  
 ; US-08-484-223B-84

Query Match 32.8%; Score 45; DB 3; Length 28;  
 Best Local Similarity 42.9%; Pred. No. 15;  
 Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 KIRALKWKNKHLKQETIAALEQ 21  
 Db 7 KVEELLSKNYHLENEVARLKK 27

RESULT 8  
 US-08-919-597-84  
 ; Sequence 84, Application US/08919597  
 ; Patent No. 6054265  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bolognesi, Dani P.  
 ; APPLICANT: Matthews, Thomas J.  
 ; APPLICANT: Wild, Carl T.  
 ; APPLICANT: Barney, Shawn O.  
 ; APPLICANT: Lambert, Dennis M.  
 ; APPLICANT: Petteway, Stephen R.  
 ; APPLICANT: Langlois, Alphonse J.  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION  
 ; OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
 ; TITLE OF INVENTION: TRANSMISSION  
 ; NUMBER OF SEQUENCES: 273  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/919,597  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/470,896  
 ; FILING DATE: 06-JUN-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7872-020  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 84:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 28 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: peptide  
 ; US-08-919-597-84

Query Match 32.8%; Score 45; DB 3; Length 28;  
 Best Local Similarity 42.9%; Pred. No. 15;  
 Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 KIRALKWKNKHLKQETIAALEQ 21  
 Db 7 KVEELLSKNYHLENEVARLKK 27

RESULT 9  
 US-08-475-668A-84  
 ; Sequence 84, Application US/08475668A  
 ; Patent No. 6060065  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Barney, Shawn O.  
 ; APPLICANT: Lambert, Dennis M.  
 ; APPLICANT: Petteway, Stephen R.  
 ; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE  
 ; FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS  
 ; TITLE OF INVENTION: TRANSMISSION  
 ; NUMBER OF SEQUENCES: 211  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds LLP  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/475,668A  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7872-026  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEFAX: (212) 869-9741/8864  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 84:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 28 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: peptide  
 ; US-08-475-668A-84

Query Match 32.8%; Score 45; DB 3; Length 28;  
 Best Local Similarity 42.9%; Pred. No. 15;  
 Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 KIRALKWKNKHLKQETIAALEQ 21  
 Db 7 KVEELLSKNYHLENEVARLKK 27



Thu Apr 28 07:26:22 2005

```

; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-485-264A-84

Query Match 32.8%; Score 45; DB 3; Length 28;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KIRALKWKNKHLKQETAALEQ 21
Db 7 KVEELLSKNYHLENEVARLKK 27

RESULT 13
US-08-474-349A-84
; Sequence 84, Application US/08474349A
; Patent No. 6333395
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway Jr., Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
; TITLE OF INVENTION: VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 517
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,349A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:

```

```

; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-474-349A-84

Query Match 32.8%; Score 45; DB 3; Length 28;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KIRALKWKNKHLKQETAALEQ 21
Db 7 KVEELLSKNYHLENEVARLKK 27

RESULT 14
US-08-255-208A-20
; Sequence 20, Application US/08255208A
; Patent No. 6440656
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway Jr., Stephen R.
; TITLE OF INVENTION: SYNTHETIC PEPTIDE INHIBITORS OF HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/255,208A
; FILING DATE: 07-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-255-208A-20

Query Match 32.8%; Score 45; DB 4; Length 28;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KIRALKWKNKHLKQETAALEQ 21
Db 7 KVEELLSKNYHLENEVARLKK 27

RESULT 15
US-08-470-896-84
; Sequence 84, Application US/08470896
; Patent No. 6479055
; GENERAL INFORMATION:

```



```

; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteaway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,896
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-470-896-84

Query Match      32.8%; Score 45; DB 4; Length 28;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 KIRALKWKNHKLKQEIQAALQ 21
      |: | | | | | | | | | |
DB      7 KVEELLSKNYHLENEVARLKK 27

```

Search completed: April 27, 2005, 15:19:43  
Job time : 30.5 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2005, 15:02:43 ; Search time 89.5 Seconds  
(without alignments)  
104.113 Million cell updates/sec

Title: US-10-088-417A-4

Perfect score: 137  
Sequence: 1 KIRALKWNAHLKQIEIAALEQ 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 334731

Minimum DB seq length: 0  
Maximum DB seq length: 28

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	81	59.1	28	14	US-10-360-053-23
2	59	43.1	28	14	US-10-360-053-21
3	54	39.4	28	14	US-10-360-053-22
4	51	37.2	24	14	US-10-338-083-19
5	51	37.2	24	16	US-10-611-399-19
6	51	37.2	24	17	US-10-794-751-19
7	50	36.5	23	10	US-09-259-658-59
8	47	34.3	21	9	US-09-834-759-534
9	47	34.3	21	13	US-10-007-805-534
10	47	34.3	21	13	US-10-007-805-587
11	47	34.3	21	14	US-10-076-622-534
12	47	34.3	21	14	US-10-076-622-587
13	47	34.3	21	14	US-10-124-805-534

14	47	34.3	21	14	US-10-124-805-587	Sequence 587, Appl
15	45	32.8	26	16	US-10-630-926-8	Sequence 8, Appli
16	45	32.8	28	14	US-10-080-608A-1	Sequence 1, Appli
17	45	32.8	28	15	US-10-370-685-90	Sequence 90, Appl
18	45	32.8	28	15	US-10-267-682-84	Sequence 84, Appl
19	45	32.8	28	15	US-10-267-748-84	Sequence 84, Appl
20	44	32.1	24	10	US-09-491-614-30	Sequence 30, Appl
21	43	31.4	20	14	US-10-076-622-596	Sequence 596, Appl
22	43	31.4	20	14	US-10-124-805-596	Sequence 596, Appl
23	42	30.7	26	14	US-10-059-720-37	Sequence 37, Appl
24	41	28.9	27	10	US-09-882-774-7	Sequence 7, Appli
25	39	28.5	17	14	US-10-210-023-111	Sequence 111, Appl
26	39	28.5	28	9	US-09-320-907B-6	Sequence 6, Appli
27	39	28.5	28	15	US-10-447-292-6	Sequence 6, Appli
28	38	27.7	25	14	US-10-059-720-2	Sequence 2, Appli
29	38	27.7	26	14	US-10-059-720-3	Sequence 3, Appli
30	38	27.7	28	14	US-10-023-066A-70	Sequence 70, Appl
31	38	27.7	28	16	US-10-804-678-70	Sequence 70, Appl
32	36	26.3	19	14	US-10-059-720-36	Sequence 36, Appl
33	36	26.3	24	14	US-10-059-720-1	Sequence 1, Appli
34	36	26.3	24	14	US-10-338-083-20	Sequence 20, Appl
35	36	26.3	24	16	US-10-611-399-20	Sequence 20, Appl
36	36	26.3	24	17	US-10-794-751-20	Sequence 20, Appl
37	36	26.3	28	14	US-10-023-066A-36	Sequence 36, Appl
38	36	26.3	28	14	US-10-023-066A-40	Sequence 40, Appl
39	36	26.3	28	14	US-10-023-066A-56	Sequence 56, Appl
40	36	26.3	28	14	US-10-023-066A-67	Sequence 67, Appl
41	36	26.3	28	16	US-10-804-678-36	Sequence 36, Appl
42	36	26.3	28	16	US-10-804-678-40	Sequence 40, Appl
43	36	26.3	28	16	US-10-804-678-56	Sequence 56, Appl
44	36	26.3	28	16	US-10-804-678-67	Sequence 67, Appl
45	35	25.5	20	14	US-10-225-567A-1341	Sequence 1341, Ap

ALIGNMENTS

RESULT 1

US-10-360-053-23  
; Sequence 23, Application US/10360053  
; Publication No. US20030170230A1  
; GENERAL INFORMATION:  
; APPLICANT: Caterer, Nigel  
; APPLICANT: Uttenhal, Lars O  
; APPLICANT: Nielsen, Rasmus W  
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibi  
; TITLE OF INVENTION: Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Regi  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: IMX-0028  
; CURRENT APPLICATION NUMBER: US/10/360,053  
; CURRENT FILING DATE: 2003-02-05  
; PRIOR APPLICATION NUMBER: US60/354,376  
; PRIOR FILING DATE: 2002-02-05  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 23  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-360-053-23

Query Match 59.1%; Score 81; DB 14; Length 28;  
Best Local Similarity 64.0%; Pred. No. 0.0084; Mismatches 3; Indels 0; Gaps 0;  
Matches 16; Conservative 6;

QY 4 ALKWKNAHLKQIEIAALEQ 28

DB 1 AIKYKQAIAKNEIAIKQIEIAIEQ 25

RESULT 2

Thu Apr 28 07:26:23 2005

```
US-10-360-053-21
; Sequence 21, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Uttenenthal, Lars O
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibiotic Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-21

Query Match 43.1%; Score 59; DB 14; Length 28;
Best Local Similarity 40.0%; Pred. No. 0.69;
Matches 10; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 4 ALAKWNAHLKQETAALEQETAALEQ 28
DB 1 AIEYQAAIKETAAIKNDIAIKE 25

RESULT 3
US-10-360-053-22
; Sequence 22, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Uttenenthal, Lars O
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibiotic Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-22

Query Match 39.4%; Score 54; DB 14; Length 28;
Best Local Similarity 50.0%; Pred. No. 3.2;
Matches 12; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 KIRALKWNAHLKQETAALEQETAA 24
DB 5 KIAAIEEKIAQIEEIAAQEEKIA 28

RESULT 4
US-10-338-083-19
; Sequence 19, Application US/10338083
; Publication No. US20030166559A1
; GENERAL INFORMATION:
; APPLICANT: Desjarlais, John R.
; APPLICANT: Tansey, Malu Lourdes G.
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: Dominant Negative Proteins and Methods Thereof
; FILE REFERENCE: A-71273-2
; CURRENT APPLICATION NUMBER: US/10/338,083
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/345,805
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/373,453
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: coiled-coil motif
US-10-611-399-19

Query Match 37.2%; Score 51; DB 16; Length 24;
Best Local Similarity 55.6%; Pred. No. 6.7;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 10 AHLKQETAALEQETAALE 27
DB 1 AALESEVSALESEVASLE 18

RESULT 5
US-10-611-399-19
; Sequence 19, Application US/10611399
; Publication No. US20040170602A1
; GENERAL INFORMATION:
; APPLICANT: Desjarlais, John R.
; APPLICANT: Tansey, Malu Lourdes G.
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: DOMINANT NEGATIVE PROTEINS AND METHODS THEREOF
; FILE REFERENCE: A-71273-3
; CURRENT APPLICATION NUMBER: US/10/611,399
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 10/338,083
; PRIOR FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US 60/345,805
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/373,453
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: coiled-coil motif
US-10-611-399-19

Query Match 37.2%; Score 51; DB 16; Length 24;
Best Local Similarity 55.6%; Pred. No. 6.7;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 10 AHLKQETAALEQETAALE 27
DB 1 AALESEVSALESEVASLE 18

RESULT 6
US-10-794-751-19
; Sequence 19, Application US/10794751
; Publication No. US20050048626A1
; GENERAL INFORMATION:
```

APPLICANT: Desjarlais, John R.  
APPLICANT: Thomason, Adam Read  
APPLICANT: Zhukovsky, Eugene Alexander  
TITLE OF INVENTION: BAPF VARIANTS AND METHODS THEREOF  
FILE REFERENCE: A-72175-1  
CURRENT APPLICATION NUMBER: US/10/794,751  
CURRENT FILING DATE: 2004-03-05  
PRIOR APPLICATION NUMBER: US 10/338,083  
PRIOR FILING DATE: 2003-01-06  
PRIOR APPLICATION NUMBER: US 60/452,707  
PRIOR FILING DATE: 2003-03-07  
PRIOR APPLICATION NUMBER: US 60/482,081  
PRIOR FILING DATE: 2003-06-23  
PRIOR APPLICATION NUMBER: US 60/523,880  
PRIOR FILING DATE: 2003-11-20  
PRIOR APPLICATION NUMBER: US 60/528,104  
PRIOR FILING DATE: 2003-12-08  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patent in version 3.3  
SEQ ID NO 19  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: coiled-coil motif  
US-10-794-751-19

Query Match 37.2%; Score 51; DB 17; Length 24;  
Best Local Similarity 55.6%; Pred. No. 6.7; Indels 3; Mismatches 5; Gaps 0;  
Matches 10; Conservative 5; Indels 3; Mismatches 5; Gaps 0;

QY 10 AHLKQEIATALEQEIATALE 27  
DB 1 AALESEVSALAESEVSL 18

RESULT 7  
US-09-259-658-59  
Sequence 59, Application US/09259658  
Publication No. US20030032054A1  
GENERAL INFORMATION:  
APPLICANT: Colyer  
APPLICANT: Craig  
APPLICANT: Maschio  
APPLICANT: Mezna  
TITLE OF INVENTION: Compositions And Methods For Monitoring The  
Modification State Of A Pair Of Polypeptides  
FILE REFERENCE: Colyer 4256/79245  
CURRENT APPLICATION NUMBER: US/09/259,658  
CURRENT FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 59  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
peptide used to assay geranylgeranyl transferase  
OTHER INFORMATION: activity.  
US-09-259-658-59

Query Match 36.5%; Score 50; DB 10; Length 23;  
Best Local Similarity 56.5%; Pred. No. 8.7; Indels 7; Mismatches 3; Gaps 0;  
Matches 13; Conservative 3; Indels 7; Mismatches 3; Gaps 0;

QY 2 IRALKKWAHLKQEIATALEQEIATA 24  
DB 1 IAQLKQKIAQLKQKIAQLKQKIA 23

RESULT 8  
US-09-834-759-534

Sequence 534, Application US/09834759  
Publication No. US20020085998A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yuqiu  
APPLICANT: Dillon, Devin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
DIAGNOSIS OF BREAST CANCER  
FILE REFERENCE: 210121.470C9  
CURRENT APPLICATION NUMBER: US/09/834,759  
CURRENT FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 547  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 534  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-834-759-534

Query Match 34.3%; Score 47; DB 9; Length 21;  
Best Local Similarity 57.9%; Pred. No. 20; Indels 5; Mismatches 3; Gaps 0;  
Matches 11; Conservative 3; Indels 5; Mismatches 3; Gaps 0;

QY 9 NAHLKQEIATALEQEIATALE 27  
DB 1 NCMLKKEIATMLKLEIATLK 19

RESULT 9  
US-10-007-805-534  
Sequence 534, Application US/10007805  
Publication No. US20020150581A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yuqiu  
APPLICANT: Dillon, Devin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedwick, Thomas S.  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Durham, Margarita  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
AND DIAGNOSIS OF BREAST CANCER  
FILE REFERENCE: 210121.470C10  
CURRENT APPLICATION NUMBER: US/10/007,805  
CURRENT FILING DATE: 2001-12-07  
NUMBER OF SEQ ID NOS: 593  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 534  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-007-805-534

Query Match 34.3%; Score 47; DB 13; Length 21;  
Best Local Similarity 57.9%; Pred. No. 20; Indels 5; Mismatches 3; Gaps 0;  
Matches 11; Conservative 3; Indels 5; Mismatches 3; Gaps 0;

QY 9 NAHLKQEIATALEQEIATALE 27  
DB 1 NCMLKKEIATMLKLEIATLK 19

RESULT 10  
US-10-007-805-587  
Sequence 587, Application US/10007805

```

; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 587
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-587

Query Match      34.3%; Score 47; DB 13; Length 21;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      9 NAHLKQEIATLAEQEIATL 27
Db      1 NCMLKKEIATLAEQEIATL 19

RESULT 11
US-10-076-622-534
; Sequence 534, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 534
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-534

Query Match      34.3%; Score 47; DB 14; Length 21;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      9 NAHLKQEIATLAEQEIATL 27
Db      1 NCMLKKEIATLAEQEIATL 19

RESULT 12
US-10-076-622-587
; Sequence 587, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.

```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 587
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-587

Query Match      34.3%; Score 47; DB 14; Length 21;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      9 NAHLKQEIATLAEQEIATL 27
Db      1 NCMLKKEIATLAEQEIATL 19

RESULT 13
US-10-124-805-534
; Sequence 534, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 534
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-534

Query Match      34.3%; Score 47; DB 14; Length 21;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      9 NAHLKQEIATLAEQEIATL 27
Db      1 NCMLKKEIATLAEQEIATL 19

RESULT 14
US-10-124-805-587
; Sequence 587, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 587
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-587

```

Matches	11;	Conservative	3;	Mismatches	5;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

• •

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

• •

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

• •

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

COUNTRY: US  
EID: 20001

FILING DATE: 28-APR-1997  
ATTORNEY/AGENT INFORMATION:

TELEFAX: 202-737-3528

118-10-520-926-8  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Matches	9: Conservative	5: Mismatches	7: Indels	0: Gaps
Matches	9: Conservative	5: Mismatches	7: Indels	0: Gaps

[illegible]

Job time : 90.5 sec

**THIS PAGE LEFT BLANK**